Applied Acoustics 75 (2014) 1–9

Contents lists available at SciVerse ScienceDirect

Applied Acoustics

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

An intelligent approach for engine fault diagnosis based on Hilbert–Huang transform and support vector machine

Y.S. Wang *, Q.H. Ma, Q. Zhu, X.T. Liu, L.H. Zhao

School of Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, PR China

ARTICLE INFO

Article history:
Received 23 October 2012
Received in revised form 30 June 2013
Accepted 1 July 2013
Available online 30 July 2013

Keywords:
Engine fault diagnosis
Noise feature vector
Wavelet packet denoising
Hilbert–Huang transform (HHT)
Support vector machine (SVM)

ABSTRACT

Based on the techniques of Hilbert–Huang transform (HHT) and support vector machine (SVM), a noise-based intelligent method for engine fault diagnosis (EFD), so-called HHT–SVM model, is developed in this paper. The noises of a sample engine under normal and several fault states are first measured and denoised by using the wavelet packet threshold method to initially lower the noise level with negligible signal distortion. To extract fault features of the engine, then, the HHT is selected and applied to the measured noise signals. A nine-dimensional vector, which consists of seven intrinsic mode functions (IMFs) from the empirical mode decomposition (EMD), maximum value of HHT marginal spectrum and its corresponding frequency component, is specified to represent each engine fault feature. Finally, an optimal SVM model is established and trained for engine failure classification by using the fault feature vectors of the noise signals. Cross-validation results show that the proposed noise-based HHT–SVM method is accurate and effective for engine fault diagnosis. Due to outstanding time–frequency characteristics and pattern recognition capacity of the HHT and SVM, the newly proposed HHT–SVM can be used to deal with both the stationary and nonstationary signals, and even the transient ones. In the view of applications, the HHT–SVM technique may be suggested not only to detect the abnormal states of vehicle engines, but also to be extended to other fields for failure diagnosis in engineering.

1. Introduction

Engine is the power source and a core component of vehicles, whose performance directly affects vehicle's safety and reliability. Due to the complex structure and working conditions of the engines, the engine fault generally constitute about 40% of the failures of whole vehicle. Therefore, it is of great significance to develop the approaches for quick and accurate condition monitoring and fault diagnosis of the engines. In the past few decades, research on the engine fault diagnosis (EFD) technology has become an active area in the field of vehicle engineering [1–4]. From 1970s, the EFD technologies have been made rapid development. The corresponding equipment, such as the diagnostic instruments with microprocessor, on-board diagnostic systems, etc., were developed and used in engineering.

According to the source of failure information, the EFD approaches can be generally classified into four kinds: (a) engine performance detection, (b) lubricating oil analysis, (c) vibration-based and (d) noise-based diagnostic methods. Since the output characteristic parameters of an engine are often not sensitive to the failures in the early stage, the performance detection method is effective only in case of the engine faults are very serious. The oil analysis method, which can be used to specify some engine faults by analyzing the wear particles in engine lubricating oil, is difficult to be applied in engineering, due to the high equipment costs and time-consuming in its detection procedure. As it is known that the vibration and noise signals of an engine contain substantially useful information (including faults), which may accurately reflect the running state of engines. The vibration- and noise-based EFD methods normally extract the failure features from the vibration or noise signals of a running engine and make a decision of diagnostic results by using pattern recognition algorithms [5,6]. The latter two methods have some outstanding characteristics, such as fast diagnostic speed and high precision and accurate positioning of fault sources. Contrast to the vibration-based ones, the noise-based techniques with the advantages of non-contact and non-disintegration have been extensively paid attention in recent years [7,8].

There have been two key points in using the noise-based EFD methods. One is an appropriate selection of the feature extraction approaches; and the other is a reasonable determination of the pattern recognition algorithms. Under the actual working conditions, engine noises are typically time-varying and nonstationary. To

* Corresponding author. Address: School of Automobile Engineering, Shanghai University of Engineering Science, No. 333, Longteng Road, Songjiang District, Shanghai 201620, PR China. Tel.: +86 21 67791149; fax: +86 21 67791152.
E-mail address: jzwbt@163.com (Y.S. Wang).

0003-682X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.apacoust.2013.07.001

extract features from the nonstationary noise signals, some time–frequency techniques, such as short-time Fourier transform (STFT), wavelet-based methods and Wigner–Ville distributions (WVD), have been frequently mentioned in literatures [3,9–11]. The STFT uses a standard Fourier transform over several types of window. The wavelet-based techniques can apply a mother wavelet function with either discrete or continuous scales to recuperate the fixed time–frequency resolution issue inherent in the STFT. The WVD with excessive transformation durations can give a better resolution than the STFT, but suffers from cross-term interference and produces results with coarser granularity than the wavelet-based methods do. A time–frequency representation called Hilbert–Huang transform (HHT), which is combined by the empirical mode decomposition (EMD) and Hilbert transform, was proposed in the late 1990s [12] and continuously discussed by researchers for its applications in engineering [13–15]. A comparison of the STFT, wavelet transforms, WVD, pseudo-WVD and HHT suggested that, the HHT with a high time–frequency resolution can clearly describe the rules of the frequency compositions changing with time, is a good approach for feature extraction in nonstationary signal processing [16]. Based on the above findings, therefore, the HHT approach is directly adopted for feature extraction of the engine failures in the noise-based EFD in this paper.

In view of the pattern recognition and classification of sound signals, some techniques in common use have been investigated and compared [17]. The results showed that: the self-organizing maps and learning vector quantization methods are complementary each other, while long-term statistics cannot be applied in combination with nonstationary feature extraction techniques. The Gaussian mixture model is effective and powerful for unsupervised classification of the musical signals. In addition to the analytical and knowledge-based models, application of data-based models has established a firm position in the fault diagnostics during the last decade. Some intelligent approaches, such as the neural network (NN), genetic algorithm (GA) and fuzzy logic algorithm (FLA), were introduced into and widely discussed on the purposes of machine fault diagnosis [4,6,8]. In 1989, a NN was established and applied in fault diagnosis and compared with the professional systems based on knowledge [1], which suggested that the NN-based methods are suitable for online processing of fault diagnosis issues. A method based on the FLA was developed and performed for EFD, which achieved an accuracy of fault diagnosis up to 90% [2]. As a novel machine learning method, the support vector machine (SVM) was introduced in early 1990s, based on the statistical learning theory, SVMs representing an extension to nonlinear models of the generalized portrait algorithm has been successfully applied to numerous classification and pattern recognition problems, such as texture categorization, machine failure recognition and chemometrics [18–21]. The SVM-based classifier is established to minimize the structural misclassification risk, whereas conventional classification methods often apply a minimization of the empirical risk. Thus, the SVM is claimed to lead some enhanced generalization properties, which may be very useful in fault diagnostics [22–24]. However, the SVMs have not yet been widely investigated for engine fault diagnostics, especially for the noise-based EFDs. Thus, this paper attempts to use the SVM technique to project the input noise features to the output fault patterns of the engines.

From the above discussion, it may be seen that the EFD methods have been developing from the traditional to the intelligent ways. For the noise-based EFDs, specially, the approaches for noise feature extraction and fault pattern classification need to be carefully selected and designed. Based on some measured engine noises, in this presented work, a new noise-based EFD technique called HHT–SVM is developed by combining the HHT and SVM algorithms and validated by tests. The scheme of this research is shown in Fig. 1. In view of the application, the HHT–SVM model can be directly used to detect faults and failures of the sample engine, and may be possibly extended to other sound-related fields for failure diagnosis in engineering, such as speech recognition and machine failure diagnosis and clinical diagnosis in medical treatment.

2. Theory background

2.1. Hilbert–Huang transform

The Hilbert–Huang transform (HHT), which is combined by the empirical mode decomposition (EMD) and the Hilbert spectral analysis, is an adaptive time–frequency analysis method designed specifically for analyzing data from nonlinear and nonstationary processes. The kernel part of the HHT is the EMD approach with which any complicated signal can be decomposed into intrinsic mode functions (IMFs). This decomposition method operating in the time domain is adaptive and highly efficient. With the Hilbert transform, the IMFs yield instantaneous frequencies as functions of time that give sharp identifications of signal components. Since the HHT is based on the local characteristic time scale of the signal, it is applicable to nonlinear and non-stationary processes.

To obtain the instantaneous frequency characteristics of a signal, in the HHT, the IMFs need to guarantee a well-behaved Hilbert transform. Thus, they are defined as functions having the same (or differing at most by one) numbers of zero-crossing and extrema, and also having symmetric envelopes (with respect to time axis) defined by the local maxima and minima, respectively. In order to extract the IMFs of a complex signal in engineering, the EMD needs to be performed, which was originally motivated from the assumption that any data consists of different simple intrinsic mode oscillations, i.e., IMFs. After EMD, a signal $x(t)$ can be expressed as,

$$x(t) = \sum_{i=1}^{n} c_i + r_n \tag{1}$$

where c_i is the ith decomposed IMF of the signal $x(t)$, r_n is the residual signal. Taking the Hilbert transform on both sides of Eq. (1), the Hilbert spectrum of $x(t)$, $H(\omega,t)$, may be carried out by the following equation:

$$H(\omega,t) = \mathrm{Re}\sum_{i=1}^{n} a_i(t) e^{j\int \omega_i(t)dt} \tag{2}$$

where *Re* is the operator of real part, $a_i(t)$ and $\omega_i(t)$ denote the functions of the amplitude and instantaneous frequency, respectively. Note that the residual term r_n in Eq. (1), which occupies very little energy of the signal, is ignored. Accordingly, the marginal spectrum of Hilbert–Huang transform, $h(\omega)$, can be defined by an integrated spectrum with respect to time, i.e.,

$$h(\omega) = \int_0^T H(\omega,t)dt \tag{3}$$

where T is the length of the signal $x(t)$. One may be found that, $H(\omega,t)$ can exactly describe the signal amplitude varying on a time–frequency plane; and $h(\omega)$ reflects the amplitude changing with frequency in the entire frequency range [25]. The instantaneous frequency of IMF, which is obtained from the Hilbert transform, is very well localized in the time–frequency domain and reveals important characteristics of the signal.

2.2. Support vector machine

A SVM performs classification by constructing a multi-dimensional hyperplane that optimally separates the data into two cate-

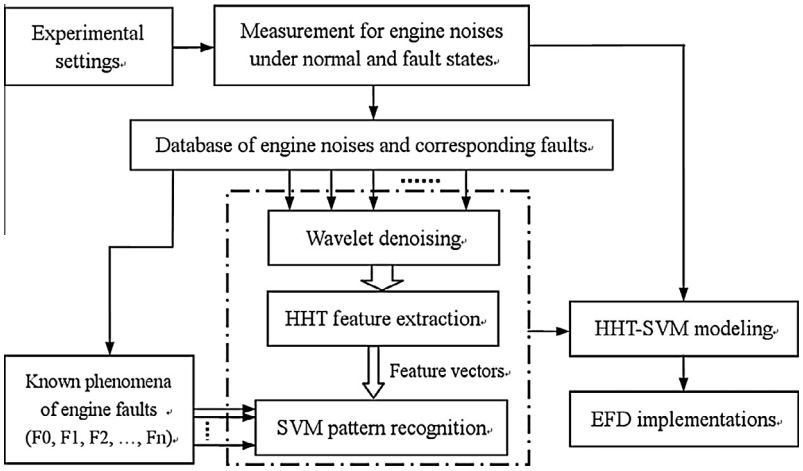

Fig. 1. The designed scheme of the HHT–SVM modeling for a noise-based EFD.

gories. The SVM modeling is to find the optimal hyperplane that separates clusters of vector in such a way that cases with one category of the target variable are on one side of the plane and cases with the other category are on the other size of the plane. The simplest way to divide two groups of data is with a straight line, flat plane or a N-dimensional hyperplane. The purpose of this paper is to classify a set of engine faults by using the extracted noise features, which actually is a nonlinear classification problem, i.e., in a linearly non-separable case. Thus, a nonlinear SVM algorithm, so-called the soft margin, is adopted in this present work.

Firstly, the data points need to be transferred into a nonlinear region. The nonlinear mapping function $\phi(x)(R^D \to \mathrm{H})$ is defined a kernel function $K(x_i,x_j) = \varphi(x_i)\varphi(x_j)$, which can map the data into a high dimensional Hilbert space H. Then, the problem for seeking an optimal separation hyperplane may be handled by solving the following optimization problem,

$$\begin{cases} \min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{L}\xi_i \\ \text{s.t.} \quad y_i(\mathrm{w}\cdot\phi(x_i)+b) \geqslant 1-\xi_i, \xi_i \geqslant 0, \quad i=1,2,\cdots,L \end{cases} \tag{4}$$

where w is normal to the hyperplane, C is a positive trade-off (cost) parameter, which defines how important it is to avoid misclassification errors, L is the training points, ξ_i is a positive slack variable, b is a constant representing the distance from the hyperplane to the origin in the Hilbert space, x_i and y_i are the input and output vectors, respectively.To meet the constraints in this minimization, the Lagrange multipliers α is allocated. And one needs to maximize:

$$L_D = \sum_{i=1}^{L}\alpha_i - \frac{1}{2}\sum_{i,j}\alpha_i H_{ij}\alpha_j = \sum_{i=1}^{L}\alpha_i - \frac{1}{2}\alpha^T H\alpha \tag{5}$$

where α_i subjects to the constraints $\sum_{i=1}^{L}\alpha_i y_i = 0$ and $0 \leqslant \alpha_i \leqslant C(\forall_i)$, $H_{ij} = y_i y_j K(x_i \cdot x_j)$ is a vector in the Hilbert space H, which is mapped from the vectors x_i and y_i. This is a convex quadratic optimization problem (QP), which can be solved by running a QP solver. Using the returned α, w may be calculated by,

$$\mathrm{w} = \sum_{i=1}^{L}\alpha_i y_i \phi(x_i) \tag{6}$$

For any support vector x_s which satisfies $\sum_{i=1}^{L}\alpha_i y_i = 0$, a variable b_s can be calculated by,

$$b_s = y_s - \sum_{m\in S}\alpha_m y_m \phi(x_m)\cdot\phi(x_s) \tag{7}$$

where S denotes a set of indices of the support vectors, which is determined by finding the indices i $(0 \leqslant \alpha_i \leqslant C)$. To define an optimal orientation of the separating hyperplane, the variable b in Eq. (4) is usually carried out by taking an average over all of the support vectors in S, i.e.,

$$b = \frac{1}{N_s}\sum_{s\in S} b_s \tag{8}$$

Substituting the variables w and b into Eq. (4), a SVM for solving a classification problem on data that is not linearly separable is created. Thus, each new point x_j may be classified by evaluating a classification function as,

$$f(x_j) = \mathrm{sgn}[w\cdot\phi(x_j)+b] = \mathrm{sgn}\left[\sum_{i=1}^{L} y_i\alpha_i K(x_i\cdot x_j)+b\right] \tag{9}$$

Some kernel mapping functions, such as the linear, polynomial, sigmoid and radial basis functions, can be considered in the SVM modeling. It has been found that the radial basis function (RBF) is very powerful in a variety of applications, thus is selected and used in this paper. The RBF known also as the Gaussian kernel is of the form,

$$K(x_i,x_j) = \exp\left(-g\|x_i - x_j\|^2\right) \tag{10}$$

where $g = 1/(2\sigma)$, σ stands for a window width.

To extend the SVM to the cases where the target variable has more than two categories, two approaches in common use are suggested: (a) "one against many" where each category is split out and all of the other categories are merged; and (b) "one against one" where a set of SVM models for two-category classifications are trained and constructed for a comprehensive recognition of the multi-category classification. Due to the small amount of computation in this paper, the method (b) with more accuracy is selected and performed. For data sets with k categories, there are $k(k-1)/2$ basic SVM models need to be established by randomly selecting two from k categories for model trainings. Following Eq. (9), $f_{ij}(x) = (w_{ij}\cdot\phi(x)) + b_{ij}$ can separate the categories i and j. A voting decision algorithm is used here. If $\mathrm{sgn}(f_{ij}(x)) = 1$, the votes of category i plus one, otherwise the votes of category j plus one. After voting by all the basic SVMs and recording the corresponding number of votes, one may finally determine a category with the largest number of votes.

3. Acquisition and pre-processing of engine noises

3.1. Data acquisitions

To develop the HHT–SVM EFD model, firstly, a database including the engine faults and their corresponding sound signals needs to be established. In this paper, based on the AJR type of GSI engine mounted on a sample car Santana 2000, the sound signals from the engine in the normal and different fault conditions were measured using the five-channel PULSE system made by the B&K Company. Considering the factors effect on the local near-field sound, such as the spatial positions of valves, pistons, rods, crankshaft, timing belt and the structure vibration of the engine, the microphone type 4189-A-021 is fixed on a point just above the center of engine cylinder head with a distance of 300 mm; and its reference direction is vertically downwards. The equipment arrangements and experimental setup for engine noise measurements are shown in Figs. 2 and 3. An engine fault generator, which may be used to generate engine faults, is placed in the rear trunk of the car. The sample engine is an inline four-cylinder four-stroke engine. The test site for noise signal acquisition is selected in a relatively opened space, which can be approximated regarded as a free field. The measured sound pressure level of background noise is 15 dB below test readings and can be ignored [26]. During the measurements, the selected acquisition parameters are: signal length, 3.0 s, sampling rate, 16,384 Hz. For a normalized scale of the measured signals, the engine is warmed up and keeping at a constant rotating speed of 2500 rpm. Seven engine working states including six typical engine faults are set by using the engine fault generator and listed in Table 1. Note that all the working states are related to engine noises. Each working state is tested twenty times. The tested data are automatically saved in the PULSE and renumbered in the format of "state no.–signal no." as shown in Table 1. We obtained 7 by 20, totally, 140 signals and saved them in a database, as well as their corresponding fault descriptions. The measured signals under each engine working state are prepared for training and testing the HHT–SVM EFD model in the following research.

3.2. Signal denoisings

In the measurement procedure, inevitably, distortion of the measured signals by certain additive noises occurred, which may possibly be some noise superpositions and interferences from the ambient background and the hardware of measurement system. Therefore, to ensure the modeling accuracy in the following text, the signals need to be denoised. Some methods for noise suppression in common use, such as least square, spectral subtraction, matching pursuit, and wavelet threshold methods have been used successfully in applications [27–29]. A shrinkage denoising method based on the discrete wavelet transform (DWT), which has been proved very effective in denoising of nonstationary vehicle noises

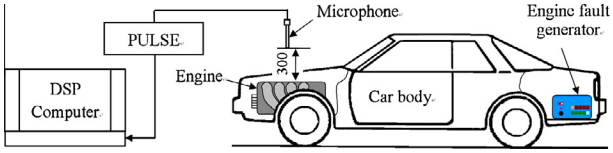

Fig. 3. Experimental setup in measurements of the emitted noises under the normal and fault engine states.

Table 1
Different settings of working states for noise measurements of the sample engine.

Engine state no.	Engine faults	Sample signal indices
1	Normal state without faults	1-1–1-20
2	A break or short in the circuit of throttle threshold sensor	2-1–2-20
3	A break or short in the circuit of ignition coil	3-1–3-20
4	A break or short in the circuit of Hall sensor G40	4-1–4-20
5	Basic set errors in the throttle control unit J338	5-1–5-20
6	A defect in the circuit of the first cylinder injector N30	6-1–6-20
7	A defect in the circuit of the third cylinder injector N32	7-1–7-20

[30], is adopted in this paper. The denoising procedure is performed in three steps: (a) signal decomposition, (b) determinations of threshold and shrinking coefficients and (c) signal reconstruction by using a self-made Matlab program. The following parameters are defined and selected: Daubechies wavelet "*db4*", 3 levels, the soft threshold signal $sign(x)\ (|x| - t)$ if $|x| > t$, otherwise is 0, where t is an universal threshold equals to $\sqrt{2\log(\text{length}(f))}$, f is frequency. As an example, the denoised signal of the noise no. 1-1 is shown in Fig. 4. It can be seen that the white noise components are well-controlled, and the denoised signal reserved the features of the original signal. This implies the DWT-based technique is effective and appropriate for the signal denoising in this paper.

4. Feature extraction of engine noises

The effectiveness of HHT in data analysis have been demonstrated by its successful application to some problems covering engineering, biomedical, financial and geophysical data [31,32]. As known that, under different fault conditions, the engine noises can show different amplitude- and phase-frequency characteristics in the frequency domain. The sound energies in some frequency bands may be obviously restrained, while enhanced in other frequency components. It should be reasonable to assume that there have certain corresponding relationships between the sound energy changes in the frequency bands of the noises and the fault

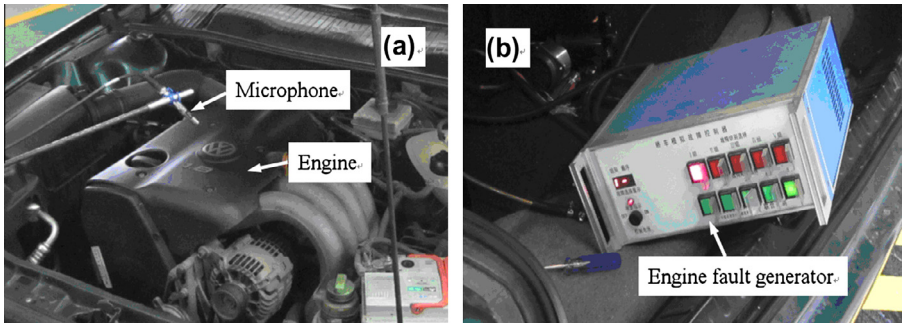

Fig. 2. Equipment arrangements for collecting engine noise signals, where (a) the microphone is mounted just above the engine; and (b) the engine fault generator is placed in the rear trunk of the car.

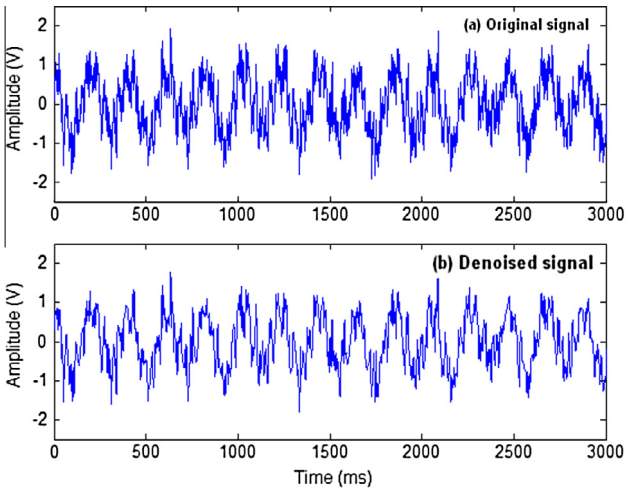

Fig. 4. An example of deniosing results of engine noise signal no. 1-1 by using the DWT-based method.

phenomena of an engine. Therefore, in this work, based on the HHT, energy pattern of the intrinsic mode function (IMF) components is considered to extract the fault features of the engine noises. The traditional IMF energy methods, such as the IMF energy ratio method, ignore the time parameter. To show more detail information of noise signals in the time and frequency domains, a time-based IMF energy moment method is proposed and performed for calculating the feature vectors in four steps.

(a) Decomposition of IMF components (IMFs). Following the procedure in Fig. 5, the IMFs of all the denoised signals are calculated by using the empirical mode decomposition (EMD). As an example, the calculated IMFs of the noise signal no. 1-1 are shown in Fig. 6. One may observed that there are nine IMFs and a residual were decomposed from the original signal.

(b) Computation of correlation coefficients. Correlation coefficient between a calculated IMF component $c_i(t)$ and its original signal $S(t)$ may be defined as,

$$\rho_{s,c_i} = \left| \frac{E[(c_i(t) - \mu_{c_i})(S(t) - \mu_s)]}{\sigma_{c_i}\sigma_s} \right| \tag{11}$$

where μ_{c_i}, σ_{c_i} and μ_S, σ_S are the mean values and standard deviations of the $c_i(t)$ and $S(t)$, respectively. ρ_{s,c_i} is the correlation coefficient $(0 \leqslant \rho \leqslant 1)$, a bigger ρ_{s,c_i} value means a greater correlation between the $c_i(t)$ and $S(t)$. One may take the major components from the IMFs and abandon the pseudo components by comparing the correlation coefficients. We select some noise signals, which involves in each fault state of the sample engine, from the prepared database and calculated the correlation coefficients. The results are listed in Table 2. A conclusion can be drawn that, regardless of what kinds of fault signals, the correlation coefficients of the IMFs c_8 and c_9 are very smaller that those from c_1 to c_7 and can be ignored. Thus, only the IMFs from the levels 1–7 are taken into account in feature extraction of engine noises in this paper.

(c) Calculation of IMF energy moments. Energy moment of the ith IMF, E_i, can be computed by the following equation,

$$E_i = \sum_{k=1}^{n} \left[(k \cdot \Delta t) \cdot |C_i(k \cdot \Delta t)|^2 \right] \tag{12}$$

where Δt is the time interval, n and k are the total number and index of sampling points, and C_i $(k{\cdot}\Delta t)$ denotes the decomposition coefficient of the ith IMF at the moment of $k{\cdot}\Delta t$.

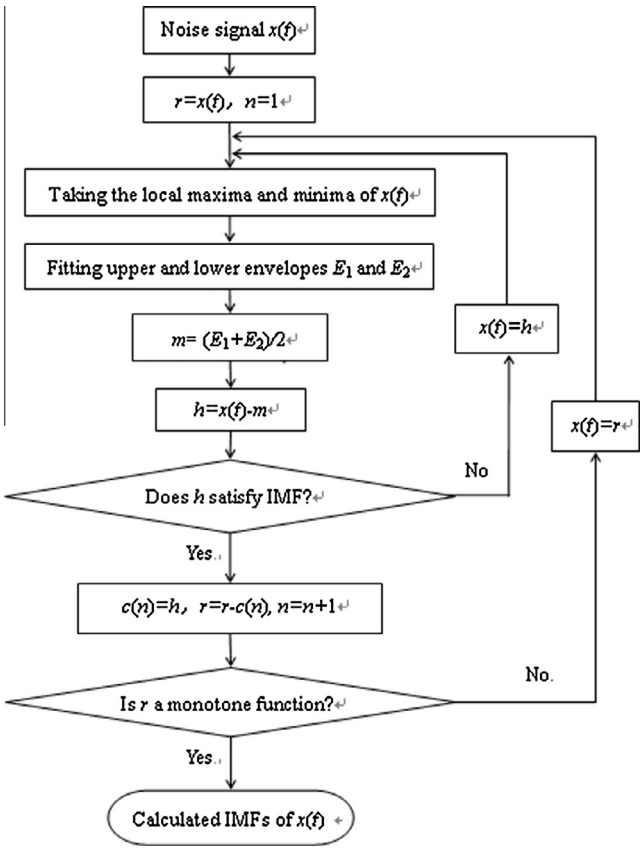

Fig. 5. The EMD procedure for IMF calculation of an engine noise signal.

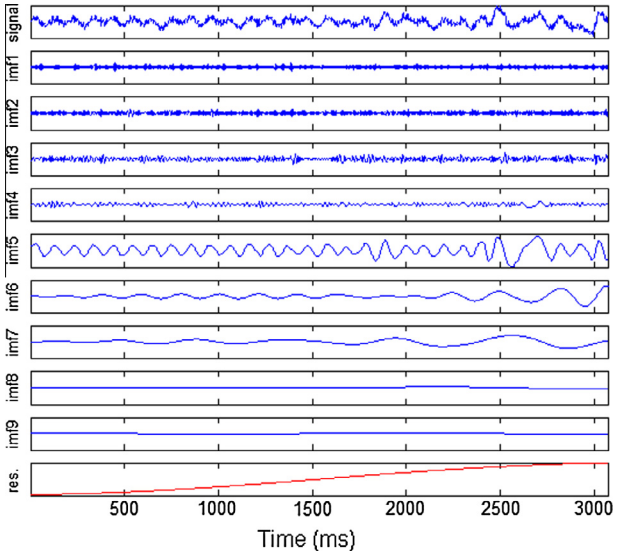

Fig. 6. An example of IMF results calculated by EMD based on the sample signal no. 1-1.

(d) Architecture of feature vectors. Taking the calculated energy moments of the former seven IMFs of each signal, E_1, E_2,..., and E_7, a seven-dimensional feature vector may be constructed as,

$$T_E = [E_1, E_2, \ldots, E_7] \tag{13}$$

Obviously, the feature vector T_E in Eq. (13) can reflect not only the energy amount of each IMF, but also the energy distributions of the IMFs changing with time, which is benefit for extracting

Table 2
Correlation coefficients between each IMF component and the original signal.

Sample no.	Correlation coefficients of the decomposed IMF components								
	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9
1-1	0.323	0.300	0.343	0.290	0.506	0.546	0.134	0.060	0.022
2-1	0.139	0.279	0.378	0.277	0.742	0.268	0.162	0.069	0.006
3-1	0.116	0.183	0.268	0.237	0.470	0.600	0.441	0.050	0.033
4-1	0.103	0.178	0.218	0.171	0.718	0.369	0.353	0.086	0.060
5-1	0.209	0.222	0.238	0.238	0.404	0.569	0.432	0.086	0.011

the fault features of the engine. In order to improve the accuracy of EFD, furthermore, marginal spectra of Hilbert–Huang transform of the engine noises, which may provide good time–frequency characteristics of a signal, are investigated in this paper. We found that, for different engine fault states, the HHT marginal spectra are different in both the maximum values and their frequency patterns. As additional feature elements, the maximum amplitude of a marginal spectrum of HHT A_0 and its corresponding frequency f_0, are introduced into the feature vector T_E. The element values of A_0 and f_0 can be directly obtained from the HHT marginal spectrum of each engine noise signal of interest. Thus, the T_E in Eq. (13) is extended to a nine-dimensional feature vector, can be rewritten as,

$$T_E = [E_1, E_2, \ldots, E_7, A_0, f_0] \tag{14}$$

The seven- and nine-dimensional feature vectors are applied and compared in the HHT–SVM modeling; and their accuracies for EFD are discussed in the following text. By using Eqs. (12) and (14), the nine-dimensional feature vectors of the measured engine noises are extracted. Due to the limitation of paper space, the feature vectors of the first five noise samples of each engine working state are shown in Figs. 7–13; and the exact feature element values of the engine noises under the normal working state are listed in Table 3. As seen from Figs. 7–13, for the noise signals measured from the same engine working state, their feature vector curves with very similar patterns are almost overlapped. For the noises under different engine working states, however, there have obvious differences in both the peak element values and the curve patterns of the extracted feature vectors. The feature differences might be used to distinguish the noise signals under different working states, thereby the corresponding engine faults. This implies that the proposed HHT-based method is effective for feature extraction of engine faults. Accuracy of the extracted feature vectors for fault diagnosis of the engine will be further discussed in section six below.

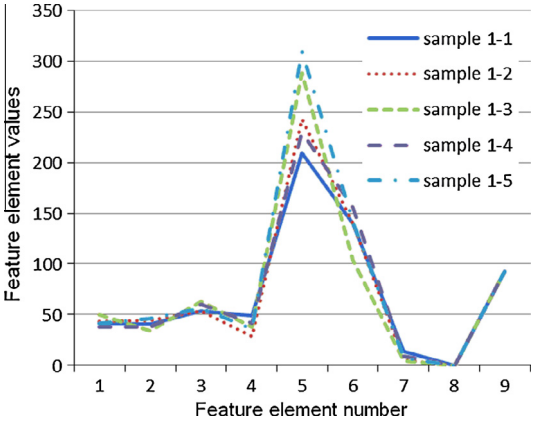

Fig. 7. Extracted feature patterns of the engine noises under the normal working state no. 1.

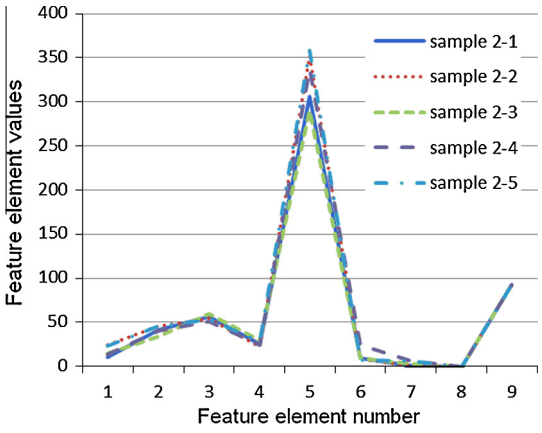

Fig. 8. Extracted feature vectors of the engine noise signals under the fault working state no. 2.

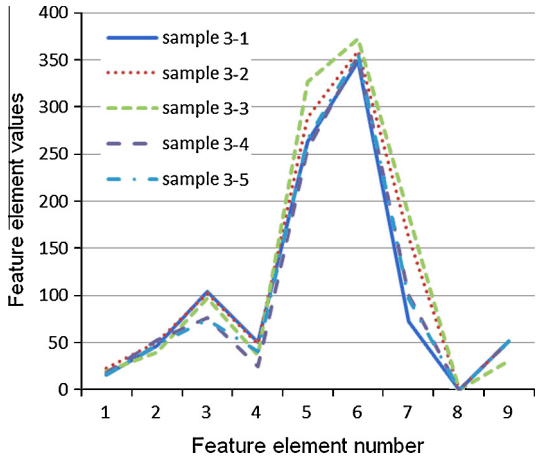

Fig. 9. Extracted feature vectors of the engine noise signals under the fault working state no. 3.

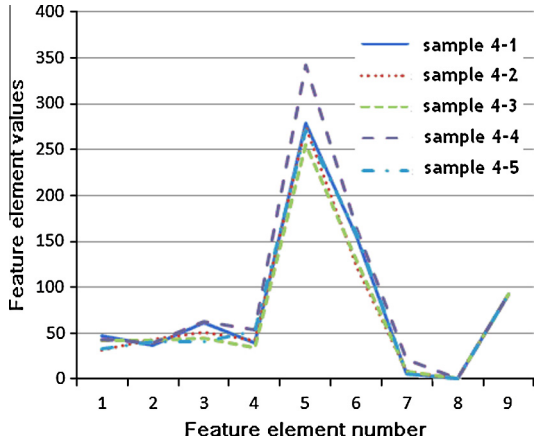

Fig. 10. Extracted feature vectors of the engine noise signals under the fault working state no. 4.

5. SVM modeling for engine fault diagnosis

The SVM diagnosis is a procedure of feature identification and classification using a well-learned SVM. A SVM-based multi-classification method with a one-to-one learning rule is introduced and applied to the EFDs in this paper, which is implemented by a Matlab program following the steps below.

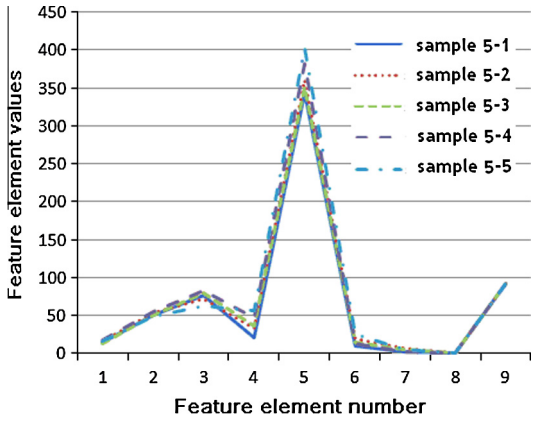

Fig. 11. Extracted feature vectors of the engine noise signals under the fault working state no. 5.

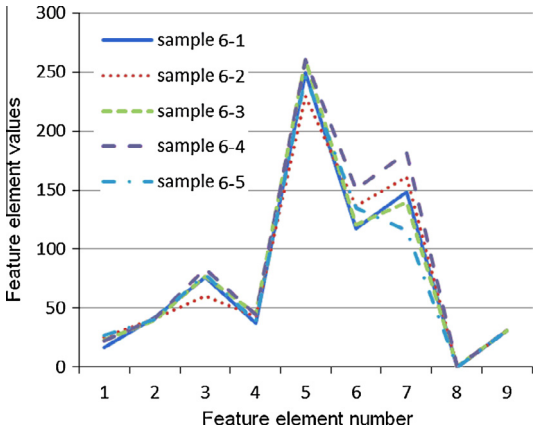

Fig. 12. Extracted feature vectors of the engine noise signals under the fault working state no. 6.

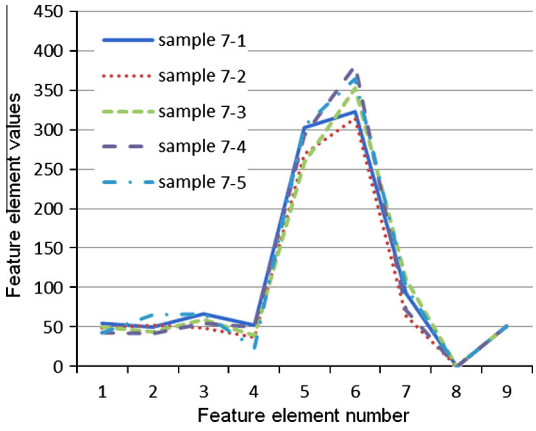

Fig. 13. Extracted feature vectors of the engine noise signals under the fault working state no. 7.

(a) Normalization of feature vectors. In order to reduce the influence of singular samples and improve EFD accuracy, the extracted feature vectors need to be normalized by using Eq. (15).

$$y = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{15}$$

Table 3
Elements of the extracted feature vectors of the engine noises under the working state no. 1.

Sample no.	Calculated element values in feature vectors of the sample signals								
	E_1	E_2	E_3	E_4	E_5	E_6	E_7	A_0	f_0
1-1	41.50	40.82	53.23	49.26	209.38	139.82	14.26	0.16	92.16
1-2	43.64	44.04	53.99	28.93	242.52	139.41	5.76	0.21	92.16
1-3	49.81	34.48	62.75	37.63	288.08	103.75	4.71	0.16	92.16
1-4	38.05	38.35	60.02	41.56	229.78	155.47	9.21	0.20	92.16
1-5	41.78	46.18	56.89	36.42	309.32	141.82	7.58	0.16	92.16

where x and y denote the extracted and normalized feature vectors, respectively.

(b) Signal grouping for SVM training and testing. The twenty noise samples of each engine state are equally divided into two sets. Ten samples are randomly selected and defined as the training set in SVM modeling, and the rest ten samples as the testing set for validation of the EFD results. Accordingly, the normalized feature vectors of the samples are naturally divided into two sets. Thus, we totally obtained two groups of samples (or feature vectors) for SVM model training and verification, respectively. Each group with seventy samples (or feature vectors) includes all of the engine working states shown in Table 1.

(c) Parameter selection in SVM modeling. As known that SVM is primarily a classier method for classification tasks by constructing hyperplanes in a multidimensional space. To construct an optimal hyperplane, the C-SVM model, which is actually performed following the calculation procedure from Eqs. (4)–(9), is selected and performed for EFDs in this paper. As mentioned before, the RBF is selected from the kernel functions in common use, because it can give localized and finite responses across the entire range of the input vectors. It can be noted that the accuracy of the C-SVM model is largely dependent on the selection of the cost parameter C in Eq. (4) and the kernel parameter g in Eq. (10). The question of the cost parameter C still remains quite an open issue. For finding optimal values of the parameters C and g, a grid search method is performed, in which the values of each parameter are tried across the search range using the specified geometric steps. To avoid over fitting, cross-validation is used to evaluate the fitting provided by each parameter value set tried during the grid search process. In view of the accuracy of EFDs, an optimal parameter combination is finally reached on the point of $C = 0.0039$ and $g = 1$, as seen in Fig. 14. As a summary, the main parameters selected in the SVM modeling are listed in Table 4. It should be mentioned that, to construct an optimal hyperplane, the SVM performs an iterative training algorithm for minimizing an error function. According to the form of error functions, the basic SVM models for classification may be classified into two types: C-SVM and nu-SVM. In this paper, a set of training and validation tests for the EFDs were carried out by using both the C-SVM and nu-SVM models and compared. The comparison results are listed in Table 5. It can be seen that the predicted EFD results from the C-SVM is more accurate than those from the nu-SVM, regardless of the dimensions of feature vector. Therefore, the C-SVM is determined and used as the basic SVM model for EFDs in this paper.

(d) SVM training and testing. Based on the determined parameters in Table 4, taking the feature vectors of some engine noise samples from the training group as inputs and their corresponding numbers of the engine states (faults) as outputs, a SVM model for EFD can be easily trained. Accordingly,

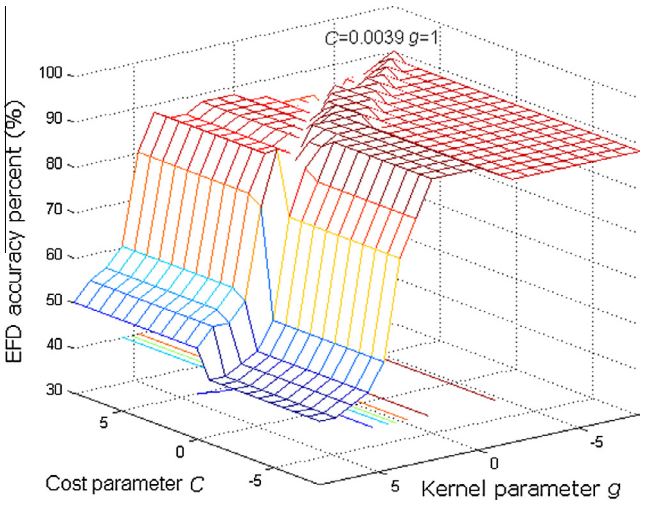

Fig. 14. An optimizing process of the cost parameter C and kernel parameter g.

Table 4
The selected and determined parameters in the SVM training in Matlab program.

Items in selection	Determined functions and parameters
Basic SVM model	C-SVM
Kernel function	Gaussian kernel (radial basis function)
Cost parameter	$C = 0.0039$
Kernel parameter Gamma	$g = 1.0$

Table 5
Comparison of the predicted EFD results by using the C-SVM and nu-SVM models.

Basic SVM model	Type of feature vector	Total samples (correctly predicted)	Accuracy rate (%)
C-SVM	Seven-dimensional	70 (56)	80.00
	Nine-dimensional	70 (64)	91.43
nu-SVM	Seven-dimensional	70 (48)	68.57
	Nine-dimensional	70 (57)	81.42

the optimal separating plane of the SVM may be determined. Similarly, one may take the feature vectors of some signals from the testing group, input them into the trained SVM model and compare the SVM outputs with the numbers of the corresponding engine fault. A percent of the matched numbers can be finally used to illustrate correctness of the trained SVM model.

6. HHT–SVM validation and application

The above mentioned seven- and nine-dimensional feature vectors are used for training the SVMs, respectively. To study the effectiveness of the SVMs in EFD, after training, the feature vectors of the sample signals in the testing group are fed to the SVM models. The SVM outputs are shown in Tables 6 and 7. Seventy testing samples involving in the normal and fault states of the engine are used in the SVM predictions. As seen from the results predicted by using the seven-dimensional feature vectors in Table 6, there have 56 states, 80.0%, are correctly predicted. In the case of the nine-dimensional feature vectors, the correction classification rate of the SVM outputs is increased to 91.43%, which is mainly attributed to the improvements in judgments of the state nos. 3, 4 and 5, as seen in Table 7. The comparisons imply that, the EFD results from the nine-dimensional feature vectors are much better than those from the seven-dimensional ones; and the feature elements A_0 and f_0 are really useful in the HHT–SVM modeling. From Tables 6 and 7, we found that Faults 4 and 5 affect so much to the EFD results, because they are sometimes confused by the normal state no. 1 and Fault 2, respectively. The reasons may be explained that, Faults 2 and 5 are the failures related to the throttle system, which cannot give a clear distinction of engine noises; and Fault 4 as a failure of the Hall sensor, which has no much effect on the engine noises, may be easily classified as the normal state no. 1. That means the noise-based HHT–SVM method proposed in this paper cannot apply to all of the engine faults. To avoid the fault selection effect on the EFDs, furthermore, a new SVM for five-state (except for the fault states 4 and 5) classification is trained and tested. The predicted accuracy of the EFDs are up to 96%, as seen in Table 8.

For more general use of the HHT–SVM model, a procedure for cross validations is performed in the presented work. In this procedure, firstly, another twenty noise signals are measured under the working states 1, 2, 3, 6 and 7 and added into the engine noise database. We randomly select ten from the saved forty signals of each working state and import them to the five-state HHT–SVM model for EFD. The cross validation is repeated many times. Two sets of the HHT–SVM results are shown in Tables 9 and 10. There

Table 6
Comparisons of the engine states and the SVM outputs by using the seven-dimensional vectors.

States (sample no.)	SVM predicted values										State no.
Normal (1-11–1-20)	1	1	1	1	1	1	1	1	1	1	1
Fault 2 (2-11–2-20)	2	2	2	5	2	2	2	2	2	5	2
Fault 3 (3-11–3-20)	3	3	3	3	5	3	5	3	3	5	3
Fault 4 (4-11–4-20)	1	4	1	4	4	1	4	4	4	1	4
Fault 5 (5-11–5-20)	5	2	5	2	2	5	5	5	2	5	5
Fault 6 (6-11–6-20)	5	6	6	6	6	6	6	6	6	6	6
Fault 7 (7-11–7-20)	3	7	7	7	7	7	7	7	7	7	7

Total samples in prediction: 70, correct predictions: 56, and accuracy rate: $A = 80.00\%$ (56/70).

Table 7
Comparisons of the engine states and the SVM outputs by using the nine-dimensional vectors.

States (sample no.)	SVM predicted values										State no.
Normal (1-11–1-20)	1	1	1	1	1	1	1	1	1	1	1
Fault 2 (2-11–2-20)	2	2	2	2	2	2	2	2	2	5	2
Fault 3 (3-11–3-20)	3	3	3	3	3	3	3	3	3	3	3
Fault 4 (4-11–4-20)	4	4	1	4	4	1	4	4	4	4	4
Fault 5 (5-11–5-20)	5	2	5	5	2	5	5	5	5	5	5
Fault 6 (6-11–6-20)	5	6	6	6	6	6	6	6	6	6	6
Fault 7 (7-11–7-20)	7	7	7	7	7	7	7	7	7	7	7

Total samples in prediction: 70, correct predictions: 64, and accuracy rate: $A = 91.43\%$ (64/70).

Table 8
The five-state HHT–SVM diagnosis results by using the nine-dimensional vectors.

States (sample no.)	SVM predicted values										State no.
Normal (1-11–1-20)	1	1	1	1	1	1	1	1	1	1	1
Fault 2 (2-11–2-20)	2	2	2	2	2	2	2	2	2	2	2
Fault 3 (3-11–3-20)	4	3	3	3	3	3	3	3	3	3	3
Fault 6 (6-11–6-20)	6	6	6	3	6	6	6	6	6	6	6
Fault 7 (7-11–7-20)	7	7	7	7	7	7	7	7	7	7	7

Total samples in prediction: 50, correct predictions: 48, and accuracy rate: $A = 96.00\%$ (48/50).

Table 9
Cross-validation results of the HHT–SVM by using random assigned samples (group 1).

States	SVM predicted values										State no.
Normal	1	1	1	1	1	1	1	1	1	1	1
Fault 2	2	2	2	2	2	2	2	2	2	2	2
Fault 3	3	3	3	3	3	3	3	3	3	4	3
Fault 6	6	6	6	6	6	6	6	6	2	2	6
Fault 7	7	7	7	7	7	7	7	7	7	7	7

Total samples in prediction: 50, correct predictions: 47, and accuracy rate: A = 94.00% (47/50).

Table 10
Cross-validation results of the HHT–SVM by using random assigned samples (group 2).

States	SVM predicted values										State no.
Normal	1	1	1	1	1	1	1	1	1	1	1
Fault 2	2	2	2	2	2	2	2	2	2	2	2
Fault 3	6	3	3	3	3	3	3	3	3	3	3
Fault 6	6	6	3	6	6	6	6	6	6	6	6
Fault 7	7	7	7	7	7	7	7	7	7	7	7

Total samples in prediction: 50, correct predictions: 48, and accuracy rate: A = 96.00% (48/50).

respectively have 47 and 48 states of the 50 noise samples are correctly recognized, corresponding to 94.0% and 96.0% of accuracy rate of the EFDs. The EFD accuracies shown in the cross validations are all above 90%, which suggest that the noise-based HHT–SVM EFD method proposed in this paper can be popularized and applied in the engineering of engine fault diagnosis.

7. Summary and conclusions

A noise-based HHT–SVM method combined by the techniques of Hilbert–Huang transform (HHT) and support vector machine (SVM) for engine fault diagnosis (EFD) is presented in this paper. Based on the HHT algorithm, the seven- and nine-dimensional vectors for feature extraction of engine noises, including the energy moments of intrinsic mode functions and the information of marginal spectra of HHT, are constructed. Some multiclass SVMs for pattern recognition in the noise-based EFDs are carefully designed, investigated and discussed. The conclusions may be drawn that, the HHT-based nine-dimensional feature vectors can be used to represent engine faults, which are much better than the seven-dimensional ones in engine fault diagnosis; the SVM with a short time-consuming in implementation is very powerful for feature recognition and classification of some noise-related faults of engines. The cross-validation results suggest that the proposed noise-based HHT–SVM method can obtain EFD accuracies more than 90%, and may be directly used in five-state EFD of the sample engine. In applications, the HHT–SVM model may be extended not only to other noise-related engine faults of vehicles, but also to other engineering fields of signal processing, such as machinery failure diagnosis, speech recognitions, etc.

Acknowledgments

This work was supported by the Project of National Natural Science Foundation of China (Grant No. 51175320), and partly supported by the Shanghai Foundation for Development of Science and Technology (Grant No. 10230501500), the Program for Professor of Special Appointment (Eastern Scholar) at the Shanghai Institutions of Higher Learning and the Fund for Talents Development by the Shanghai Municipality, China.

References

[1] Venkat V, King C. A neural network methodology for process fault diagnosis. J AICHE 1989;35(12):1993–2002.
[2] Twiddle JA, Jones NB. A high-level technique for diesel engine combustion system condition monitoring and fault diagnosis. Syst Control Eng 2002;216:125–34.
[3] Albarbar A, Gu F, Ball AD, Starr A. Acoustic monitoring of engine fuel injection based on adaptive filtering techniques. Appl Acoust 2010;71:1132–41.
[4] Si JP, Liu P, Gao ZY, Liu ZF. Fault diagnosis of engine valve clearance based on vibration signal analysis. Chin J Small Intern Combust Engine Motorcycle 2007;36(2):73–5.
[5] Li GB, Duan SL, Yu HL, Guan DL. Study on characteristic parameters of engine vibration signal based on multi-fractal. Trans CSICE 2008;26(1):87–91.
[6] Reza Tafreshi. Feature extraction using wavelet analysis with application to machine fault diagnosis. PhD thesis. University of British Columbia; 2005.
[7] Eftekharnejad Babak, Mba D. Seeded fault detection on helical gears with acoustic emission. Appl Acoust 2009;70:547–55.
[8] Lee SK, Kim SJ. Internal combustion engine sound-based fault detection and diagnosis using adaptive line enhancers. J Automob Eng 2008;222:593–605.
[9] Gold B, Morgan N. Speech and audio signal processing. New York: John Wiley & Sons; 2000.
[10] Wang YS, Lee CM, Zhang LJ. Wavelet analysis of vehicle nonstationary vibration under correlation four-wheel random excitation. Int J Automot Technol 2004;5(4):257–68.
[11] Loutas TH, Sotiriades G, Kalaitzoglou I, Kostopoulos V. Condition monitoring of a single-stage gearbox with artificially induced gear cracks utilizing on-line vibration and acoustic emission measurements. Appl Acoust 2009;70:1148–59.
[12] Huang Norden E. A new view of nonlinear waves: the Hilbert spectrum. Ann Rev Fluid Mech 1999;31(5):417–57.
[13] Huang Norden E, Wu Man-Li C, Long Steven R, Shen Samuel SP, Qu W, Gloersen P, et al. A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. Proc R Soc A: Math Phys Eng Sci 2003;459:2317–45.
[14] Yu DJ, Cheng JS, Yang Y. The Hilbert–Huang transform method for mechanical fault diagnosis. Beijing: China Science Press; 2006.
[15] Wu Z, Huang Norden E. Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 2009;1:1–41.
[16] Zhu Q, Wang YS, Shen GQ. Research and comparison of time–frequency techniques for nonstationary signals. J Comput 2012;7(4):504–8.
[17] Kohonen T. Self-organizing maps. Berlin, Germany, Printed in USA: Springer-Verlag; 1997.
[18] Xu GL, Wang XT, Xu XG. Improved bi-dimensional EMD and Hilbert spectrum for the analysis of textures. Pattern Recognit 2009;42(5):718–34.
[19] Yu DJ, Yang Y, Cheng JS. The fault diagnosis method for gear based on SVM and EMD. Chin J Mech Eng 2005;41(1):140–4.
[20] Pöyhönen S, Negrea M, Jover P, Arkkio A, Hyötyniemi H. Numerical magnetic field analysis and signal processing for fault diagnostics of electrical machines. Int J Comput Math Electr Electron Eng 2003;22(4):969–81.
[21] Ivanciuc O. Applications of support vector machines in chemistry. Rev Comput Chem 2007;23:291–400.
[22] Duda RO, Hart PE, Stork DG. Pattern classification. New York: Wiley; 2001.
[23] Zhang L, Zhou WD, Jiao LC. Scaling kernel function support vector machines. Acta Electron Sin 2002;4(4):527–9.
[24] Liu G, Liu X, Jing Q. Fault diagnosis approach based on hidden markov model and support vector machine. Chin J Mech Eng 2007;20:92–5.
[25] Huang NE, Chern CC, Huang K, Salvino LW, Long SR, Fan KL. A new spectral representation of earthquake data: Hilbert spectral analysis of station TCU129, Chi-Chi, Taiwan, 21 September 1999. Bull Seismol Soc Am 2001;91(5):1310–38.
[26] Ma DQ. Noise and vibration control manual. Beijing: China Machine Press; 2002.
[27] Bouquin RL. Enhancement of noisy speech signals: application to mobile radio communications. Speech Commun 1996;18(1):3–19.
[28] Ruiz-Reyes N, Vera-Candeas P, Curpián-Alonso J, Mata-Campos R, Cuevas-Martínez JC. New matching pursuit-based algorithm for SNR improvement in ultrasonic NDT. NDT & E Int 2005;38(6):453–8.
[29] Kim YY, Hong J-C, Lee N-Y. Frequency response function estimation via a robust wavelet de-noising method. J Sound Vib 2001;244(4):635–49.
[30] Wang YS, Lee C-M, Kim D-G, Xu Y. Sound-quality prediction for nonstationary vehicle interior noise based on wavelet pre-processing neural network model. J Sound Vib 2007;299:933–47.
[31] Huang Norden E, Shen Samuel SP. The Hilbert–Huang transform and its applications. World Scientific Publishing Company; 2005.
[32] Huang Norden E, Wu Z. A review on Hilbert–Huang transform: method and its applications to geophysical studies. Rev Geophys 2008;46:1–23.