

Research Article

A Novel Method for Mechanical Fault Diagnosis Based on Variational Mode Decomposition and Multikernel Support Vector Machine

Zhongliang Lv,^{1,2} Baoping Tang,¹ Yi Zhou,¹ and Chuande Zhou²

¹The State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing 400030, China

²College of Mechanical and Power Engineering, Chongqing University of Science and Technology, Chongqing 401331, China

Correspondence should be addressed to Baoping Tang; bptang@cqu.edu.cn

Received 4 June 2015; Accepted 25 October 2015

Academic Editor: Gyuhae Park

Copyright © 2016 Zhongliang Lv et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A novel fault diagnosis method based on variational mode decomposition (VMD) and multikernel support vector machine (MK SVM) optimized by Immune Genetic Algorithm (IGA) is proposed to accurately and adaptively diagnose mechanical faults. First, mechanical fault vibration signals are decomposed into multiple Intrinsic Mode Functions (IMFs) by VMD. Then the features in time-frequency domain are extracted from IMFs to construct the feature sets of mixed domain. Next, Semisupervised Locally Linear Embedding (SS-LLE) is adopted for fusion and dimension reduction. The feature sets with reduced dimension are inputted to the IGA optimized MK SVM for failure mode identification. Theoretical analysis demonstrates that MK SVM can approximate any multivariable function. The global optimal parameter vector of MK SVM can be rapidly identified by IGA parameter optimization. The experiments of mechanical faults show that, compared to traditional fault diagnosis models, the proposed method significantly increases the diagnosis accuracy of mechanical faults and enhances the generalization of its application.

1. Introduction

To ensure the safe and reliable operation of mechanical equipments, vibrations are usually analyzed to diagnose mechanical faults [1–3]. Accurate diagnosis helps to make reasonable maintenance decision. However, the diagnosis accuracy is generally low and manual intervention of diagnosis is usually needed [4, 5], due to the features of large rotating machinery, including multicomponent coupling vibration, strong vibration noise interference, and the instability and nonlinearity of signals, as well as the low noticeability of early fault signals.

As an adaptive method which processes signal in time-frequency domain, empirical mode decomposition (EMD) [6] can decompose the complicated and unstable signal into several nearly stable IMFs. Therefore, it has been widely applied to the diagnosis of mechanical fault [7]. For example, Loutridis [8] used EMD method for gear fault diagnosis. Cheng et al. used EMD method to diagnose bearing failures [9]. However, EMD method is essentially a binary filter bank.

The frequency domain splitting feature of EMD makes it disadvantageous in dealing with fault signals [10]. Since the band center and bandwidth of fault signal are unknown, strong interference may be introduced if the fault signal falls in the broad bands of the first component. If the fault signal is in the bands of higher-order components, important information on the feature may be missed as the signal is filtered out by the narrow band of low-order components. Dragomiretskiy and Zosso, in 2014, proposed a new adaptive signal processing method called VMD [11]. This method is able to determine the frequency center and bandwidth of each component in the process of acquiring decomposed components by iteratively searching the optimal solution of variational models, thus adaptively realizing the frequency domain split and the effective separation of each component.

In pattern recognition, support vector machine (SVM) [12] is based on Vapnik-Chervonenkis dimension theory and structural risk minimization principle. SVM finds the optimal compromise between model complexity and learning ability using limited sample information. It overcomes the

drawback that traditional machine learning models are easy to get trapped in local minima. It has enormous potential to accurately classify the faults into multiple levels. Therefore, SVM has been widely applied to all kinds of nonlinear pattern recognition problems [13]. However, in complicated cases, especially when the data are heterogeneous [14] and samples are unevenly distributed [15], or samples are in large scales [16], SVM [17] begins to lose its advantages in accomplishing the tasks. MKSVM is a new machine learning model which combines all individual kernels by weights based on traditional single kernel SVM. MKSVM inherits the generalization and learning ability of single kernel SVM. Meanwhile, it reasonably adjusts the weight of each individual kernel and improves the adaptability and robustness of single kernels [18]. However, in the fault identification process using MKSVM, the identification performance of MKSVM is directly influenced by the choice of the function parameters of individual kernels and their weights. Traditional methods, including trial and error or traversing optimization, are not only complicated in computation, but also unable to acquire global optimal solution. Therefore, the adaptive diagnosis ability of MKSVM still needs further improvements. IGA [19] finds optimal solutions by synthetically considering the information interaction between antibodies of populations. Based on genetic algorithm [20], it integrates a series of mechanisms of biological immune system, such as antigen recognition, antibody diversity, density control, and elitist strategy. It greatly helps to avoid immature convergences and meanwhile preserves the global stochastic parallel searching character of genetic algorithm at the same time [19, 21]. Taking computational efficiency, stability, and global optimality factors into consideration, IGA is used to optimize the penalty parameter, weight factor, and kernel parameters of MKSVM, in order to improve the accuracy and stability of fault diagnosis, as well as to enhance the applicability of MKSVM.

The rest of this paper is organized as follows. In Section 2, the theories of EMD and VMD are introduced. In Section 3, the theories of SVM and MKSVM are firstly reviewed. Then, the IGA optimized MKSVM (IGA-MKSVM) method is presented. In Section 4, the fault diagnosis strategy based on VMD and IGA-MKSVM is discussed. Finally, the effectiveness of the proposed method is verified by experiments.

2. VMD Method

2.1. EMD. IMF components obtained by EMD should meet the following criteria: (1) In a data sequence, the number of extreme points and the number of zero crossing points are equal, or up to a difference of 1; (2) at any data point, the average of the local maximum envelope and the local minimum envelope is 0. The basic algorithm of EMD is as follows.

Step 1. First, determine the local extremes of signal $x(t)$. Then use a cubic spline to connect all the local maximum points to form the upper envelope. After that, use another cubic spline to connect all the local minimum points to form the lower envelope. The upper and lower envelopes should enclose all

data points. The average of the upper and lower envelopes is denoted as $m(t)$. $h(t)$ refers to the new signal, which is obtained by subtracting $m(t)$ from $x(t)$:

$$h(t) = x(t) - m(t). \quad (1)$$

Repeat Step 1 by K times until $h_{1k}(t)$ becomes the basic IMF component.

Step 2. Define $c_1(t)$, $c_1(t) = h_{1k}(t)$, which is the first mode component obtained by processing the raw data. It should contain the shortest periodic component of raw signals. The residual component $r_1(t)$ can be calculated as follows:

$$r_1(t) = x(t) - c_1(t). \quad (2)$$

Step 3. Since the residual component $r_1(t)$ still contains information of long periodic components, $r_1(t)$ is still treated as new signal data. Repeat the above steps for the residual component $r_j(t)$, and the following results can be obtained:

$$\begin{aligned} r_1(t) - c_2(t) &= r_2(t), \\ r_2(t) - c_3(t) &= r_3(t), \\ &\vdots \\ r_{N-1}(t) - c_N(t) &= r_N(t). \end{aligned} \quad (3)$$

Step 4. The original signal $x(t)$ is finally decomposed into the sum of several IMF components $c_i(t)$ and a residual $r_N(t)$:

$$x(t) = \sum_{i=1}^N c_i(t) + r_N(t). \quad (4)$$

2.2. VMD. In VMD algorithm [11, 22], an intrinsic mode function is redefined as an AM-FM signal, which is expressed by

$$u_k(t) = A_k(t) \cos(\varphi_k(t)), \quad (5)$$

where $A_k(t)$ is the instantaneous amplitude of $u_k(t)$; $\omega_k(t)$ is the instantaneous frequency of $u_k(t)$ and $\omega_k(t) = \dot{\varphi}_k(t) = d\varphi(t)/dt$. $A_k(t)$ and $\omega_k(t)$ are slowly varying with respect to the phase $\varphi_k(t)$. That is, in the interval of $[t - \delta, t + \delta]$ (where $\delta = 2\pi/\dot{\varphi}_k(t)$), $u_k(t)$ can be viewed as a harmonic signal with amplitude $A_k(t)$ and frequency $\omega_k(t)$.

To obtain IMF components, VMD algorithm does not use the cycled screening stripping signal processing mode of EMD. Instead, VMD moves the signal decomposition process into the variational framework. It realizes adaptive signal decomposition by searching the optimal solution of the constrained variational model. The frequency center and bandwidth of each IMF component are updated in the iterative solving process of the variation model. The signal band is adaptively split according to the frequency domain features of the signal. Finally the narrow band IMF components are obtained.

Assuming the original signal $x(t)$ is decomposed into K IMF components, the corresponding constrained variational model is expressed as follows:

$$\begin{aligned} \min_{\{u_k, w_k\}} & \left\{ \sum_k \left\| \partial_t \left[\left(\sigma(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-jw_k t} \right\|_2^2 \right\} \\ \text{s.t.} & \sum_k u_k = x(t), \end{aligned} \quad (6)$$

where $\{u_k\} = \{u_1, \dots, u_k\}$ represents the K IMF components decomposed by VMD; $\{w_k\} = \{w_1, \dots, w_k\}$ represents the frequency centers of all IMF components.

To obtain the optimal solution of the above constrained variational problem, the following augmented Lagrange function is introduced:

$$\begin{aligned} L(\{u_k\}, \{w_k\}, \lambda) &= \alpha \sum_k \left\| \partial_t \left[\left(\sigma(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-jw_k t} \right\|_2^2 \\ &+ \left\| x(t) - \sum_k u_k(t) \right\|_2^2 \\ &+ \left\langle \lambda(t), x(t) - \sum_k u_k(t) \right\rangle, \end{aligned} \quad (7)$$

where α is a penalty factor, and λ is the Lagrange multiplier.

The optimal solution of the constrained variational model is derived by using alternating direction multiplier algorithm, which solves the saddle point of the above augmented Lagrange function. The decomposition of the original signal $x(t)$ is then obtained. The detailed implementation steps are as follows:

- (1) Initialize $\{u_k^1\}$, $\{w_k^1\}$, λ^1 , and n as 0.
- (2) $n = n + 1$; perform the entire cycle.
- (3) Perform the first inner cycle and update u_k according to:

$$u_k^{n+1} = \arg \min_{u_k} L(\{u_{i < k}^{n+1}\}, \{u_{i \geq k}^n\}, \{w_i^n\}, \lambda^n). \quad (8)$$

- (4) $k = k + 1$; repeat step (3) until $k = K$ and finish the first inner cycle.
- (5) Perform the second inner cycle and update w_k according to:

$$\begin{aligned} w_k^{n+1} &= \arg \min L(\{u_i^{n+1}\}, \{w_{i < k}^{n+1}\}, \{w_{i < k}^{n+1}\}, \{w_{i > k}^n\}, \lambda^n). \end{aligned} \quad (9)$$

- (6) $k = k + 1$; repeat step (5) until $k = K$. Finish the second inner cycle.
- (7) Update λ according to $\lambda^{n+1} = \lambda^n + \tau(f - \sum u_k^{n+1})$.
- (8) Repeat steps (2) to (7) until the criterion $\sum \|u_k^{n+1} - u_k^n\|_2^2 / \sum \|u_k^n\|_2^2 < \varepsilon$ is satisfied. Finish the entire cycle and output the results of K narrow band IMF components.

3. IGA-MKSVM

3.1. SVM. Assume two linearly separable sample sets, $\{(x_i, y_i), i = 1, \dots, n\}$, $x_i \in R^n$, $y_i \in \{+1, -1\}$. The general form of decision functions is $f(x) = \omega \cdot x + b$, and the decision surface equation is as follows:

$$\omega \cdot x + b = 0. \quad (10)$$

Normalizing the decision equation so that the samples of both classes satisfy $|f(x)| \geq 1$. The samples closest to the decision surface satisfy $f(x) = 1$. The decision surface correctly classifies all samples; that is,

$$y_i [(\omega \cdot x_i + b)] - 1 \geq 0, \quad i = 1, \dots, n. \quad (11)$$

Support vectors are samples satisfying (11) and such that $(1/2)\|\omega\|^2$ is minimum. Those samples are on the lines of H_1 and H_2 , as shown in Figure 1, where H is the optimal classification surface.

Under the constraint of (11), the problem of optimal classification surface can be expressed as follows:

$$\begin{aligned} \text{s.t.} \quad y_i [(\omega \cdot x_i + b)] - 1 &\geq 0, \quad i = 1, \dots, n, \\ \phi(\omega) &= \frac{1}{2} \|\omega\|^2. \end{aligned} \quad (12)$$

An optimal hyper surface is found such that the average classification error for the entire training sample set reaches minimum. Introducing a nonnegative relaxation factor ξ and allowing the existence of misclassified samples, the decision surface $\omega \cdot x + b = 0$ satisfies

$$y_i [(\omega \cdot x) + b] \geq 1 - \xi_i, \quad i = 1, \dots, n. \quad (13)$$

A penalty term $c \sum_{i=1}^n \xi_i$ is then added to $(1/2)\|\omega\|^2$, and the following objective function is introduced:

$$\phi(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^n \xi_i, \quad (14)$$

where c is the penalty factor.

Using Lagrange optimization method, the above optimal classification surface problem is converted into the following dual problem of convex quadratic programming optimization:

$$\begin{aligned} \max \quad & \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_i a_j y_i y_j (x_i \cdot x_j) \\ \text{s.t.} \quad & a_i \geq 0, \quad i = 1, \dots, n \\ & \sum_{i=1}^n a_i y_i = 0. \end{aligned} \quad (15)$$

By solving the above problem, the optimal decision function can be described as follows:

$$\begin{aligned} f(x) &= \text{sgn} \{(\omega \cdot x) + b\} \\ &= \text{sgn} \left\{ \sum_{i=1}^n a_i^* y_i (x_i \cdot x) + b^* \right\}. \end{aligned} \quad (16)$$

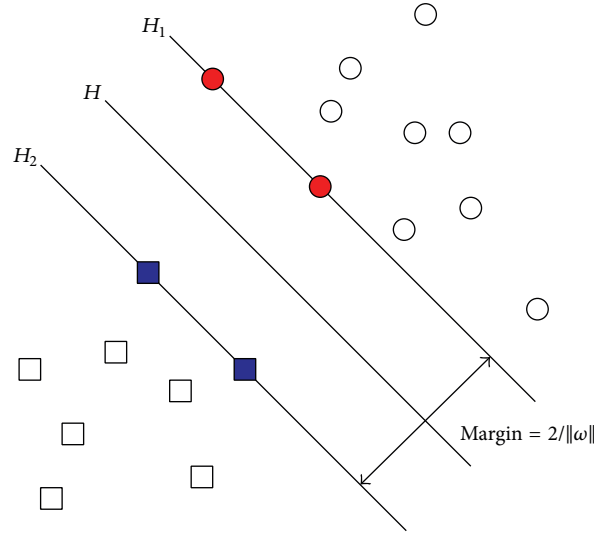


FIGURE 1: Schematic of the optimal surface.

3.2. *MKSVM*. Since different kernel functions correspond to different decision functions, the selection of kernel functions is very important in fault identification using SVM, and it directly affects the identification accuracy of SVM.

The kernel functions of SVM mainly include local kernel functions and global kernel functions. The Gaussian kernel function is a typical kernel function, which is described as follows:

$$K_{\text{RBF}}(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right). \quad (17)$$

Polynomial function is a typical global kernel function, which is described as follows:

$$K_{\text{ploy}}(x_i, x_j) = (x_i^T - x_j + 1)^d. \quad (18)$$

Local kernel functions have strong learning ability but weak generalization ability; while global kernel functions have strong generalization ability but weak learning ability. In order to achieve better learning and generalization abilities of SVM, MKSVM is constructed based on local kernels and global kernels:

$$K_{\text{mix}}(x_i, x_j) = \lambda K_{\text{RBF}}(x_i, x_j) + (1 - \lambda) K_{\text{ploy}}(x_i, x_j), \quad (19)$$

where x_i and x_j are feature vectors of input space; σ and d are kernel parameters of Gaussian kernel and polynomial kernel, respectively; λ ($0 < \lambda < 1$) is a weight factor. MKSVM combines the advantages of each single kernel and has better classification and identification performance.

3.3. *IGA-MKSVM*. IGA treats the object problem to be solved as biological invasion antigen and the feasible solution of the problem as antibody. The searching process of the

optimal solution can be viewed as the process of seeking maximum antigen affinity antibodies by biological systems. The inhibition and promotion of antibodies can ensure the diversity of antibodies in the population and improve the local searching ability of GA. Crossover and mutation of antibodies can ensure that the antibody population evolves towards the direction of high fitness and maintain the diversity of the population. The memory unit accelerates searching by constantly updating with better solutions, which improves the global searching capability of the algorithm. The flowchart of IGA algorithm is shown in Figure 2. Refer to [18, 19] for more details of IGA.

In IGA-MKSVM, IGA algorithm is used to optimize the weight factor, penalty parameter, and kernel parameter. First, an antibody gene vector X is defined, which is composed of the weight factor of MKSVM λ , penalty parameter c , and parameters σ and d of kernel function:

$$X = [\lambda, c, \sigma, d]. \quad (20)$$

To minimize the square error between the actual output and the expected output of MKSVM, the fitness function $f(x_i)$ of MKSVM is defined based on the classification accuracy of training sample $E(x_i)$:

$$f(x_i) = E(x_i). \quad (21)$$

The flowchart of IGA-MKSVM algorithm is shown in Figure 2, and the detailed steps are as follows:

- (1) Initialize the population and determine the population size, fitness threshold, and maximum iteration number. Determine the initial vector of each antibody within the ranges of penalty parameter and kernel parameters.
- (2) Compute the fitness value of each antibody according to (21) using the current antibody and training samples.

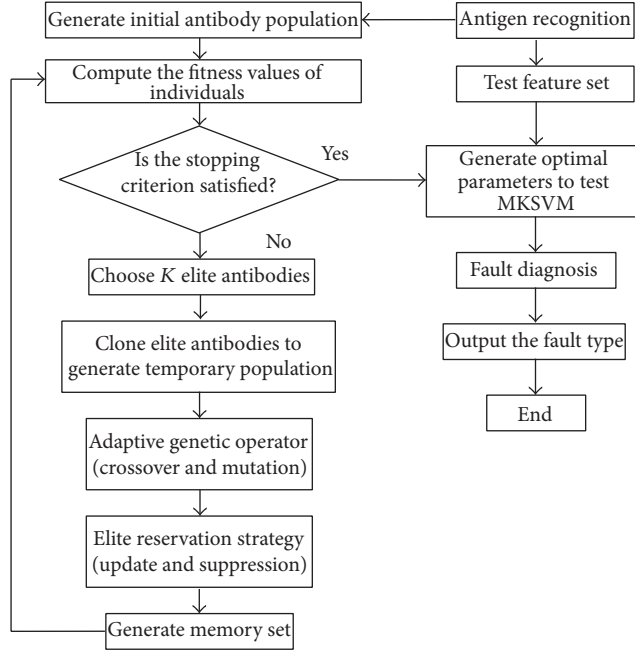


FIGURE 2: Flowchart of IGA.

- (3) For the current population, choose the antibody with the highest fitness value as the elite antibody, and save this antibody in a special variable.
- (4) If it is the first generation of antibody population, go to step (7); otherwise, go to the next step.
- (5) Determine the fitness value of each antibody. If no antibody in the current antibody generation has the same fitness value as the elite antibody, then replace the antibody having the smallest fitness value in the current antibody population by the elite antibody saved in a special variable; otherwise, go to the next step.
- (6) If the maximum fitness value in the current antibody population is larger than that of the elite antibody, then copy the antibody with the maximum fitness value into the special variable to replace the current elite antibody; otherwise, go to the next step.
- (7) According to the similarity definition, compute the density and selection probability of each antibody; perform selection and copy operations for the antibody population according to the selection probability.
- (8) Perform crossover and mutation operation for the antibody population.
- (9) Judge whether the criterion of termination is satisfied. If yes, output the results and the algorithm ends; otherwise, return to step (2) and continue the cycle.

TABLE 1: The dimensionless time domain characteristic index.

Number	Feature expression
1	$\bar{X} = \frac{1}{N} \sum_{i=1}^N x_i$
2	$X_{\text{rms}} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$
3	$\beta = \frac{1}{N} \sum_{i=1}^N x_i^4$
4	$\sigma_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2$
5	$X_r = \left[\frac{1}{N} \sum_{i=1}^N \sqrt{ x_i } \right]^2$

4. Fault Diagnostic Strategy

The flowchart of fault diagnosis based on VMD and IGA-MKSVM is shown in Figure 3.

- (1) First, decompose the obtained fault vibration signals of rotating machinery by VMD, and generate L IMF components that contain most fault information.
- (2) Extract 5 statistical features in time domain and 7 statistical features in frequency domain. These features constitute a $L \times 12$ combined domain feature set with high dimension. The indices of the 5 time domain and 7 frequency domain features are presented in Tables 1 and 2.

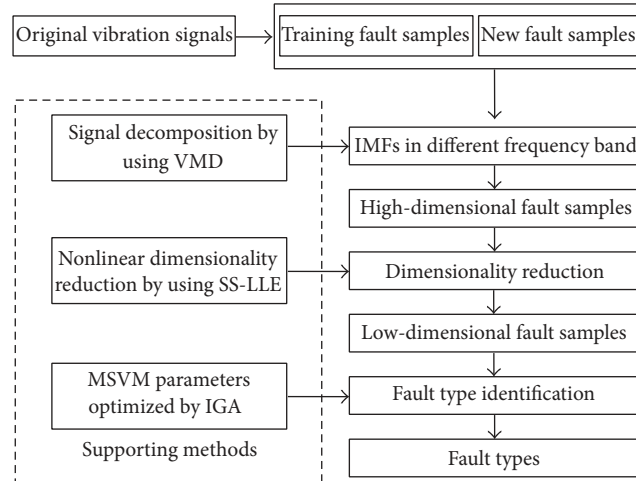


FIGURE 3: The process of fault diagnosis based on VMD and IGA-MSVM.

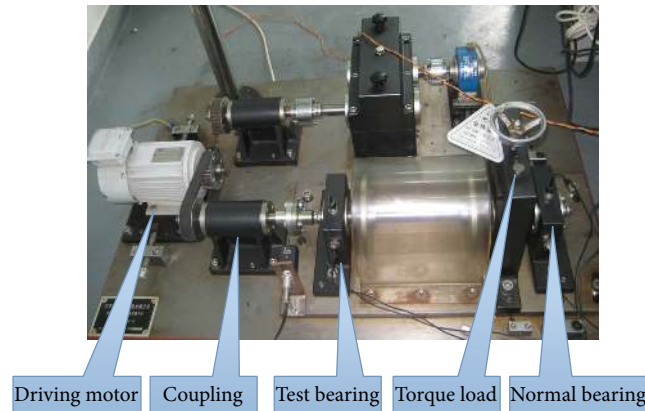


FIGURE 4: Experimental devices.

- (3) Use the SS-LLE algorithm to merge the combined high-dimensional feature set and reduce its dimension and then input the feature set to IGA-MKSVM as feature vectors.
- (4) Use training samples to train the IGA-MKSVM and obtain the optimal weight factor λ , penalty parameter c , and kernel function parameters σ and d . Then, use the optimal parameter vector to reconstruct the fault diagnosis model and identify failure modes of test samples. Finally, output the diagnostic results.

5. Examples of Application

5.1. Experiment Set-Up and Signal Acquisition. The experimental test system is composed of a speed motor, a driving belt, a coupling, a test bearing, a magnetic brake, an acceleration sensor, and a signal record analyzer. The bearing model is N205. The experimental devices are shown in Figure 4.

Three N205 bearings were used to simulate different damages, including the outer ring damage, rolling damage,

and inner ring damage. Bearing damages were implemented by processing a slot in different parts of bearings with a laser cutting machine. The width of the slot is 0.3 mm, and the depth is 0.1 mm. Different bearing faults are shown in Figure 5. Four modes of bearing vibration signals were collected. They are (I) normal, (II) inner ring damage, (III) outer ring damage, and (IV) rolling body damage. The main experimental parameters are tabulated in Table 3.

ICP accelerometers and data acquisition equipment DP/INV306U were used to collect the vibration signals of bearings. The sampling frequency is 10 kHz. In the input terminal, a low pass filter was used for antialiasing. For each operating status, 120 sample groups were collected, with a total of 10000 points for each sample group. Thirty sample groups were randomly selected as training samples, and the other 90 sample groups were test samples. Figure 6 shows the time domain waveforms of the four modes.

VMD decomposition was performed for the four modes of vibration signals. The decomposition results of VMD and EMD for normal vibration signals of bearings are shown in

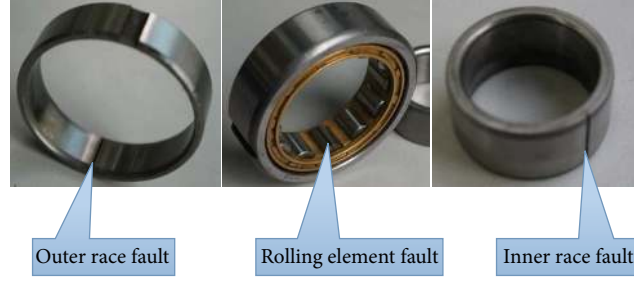


FIGURE 5: Rolling bearing fault.

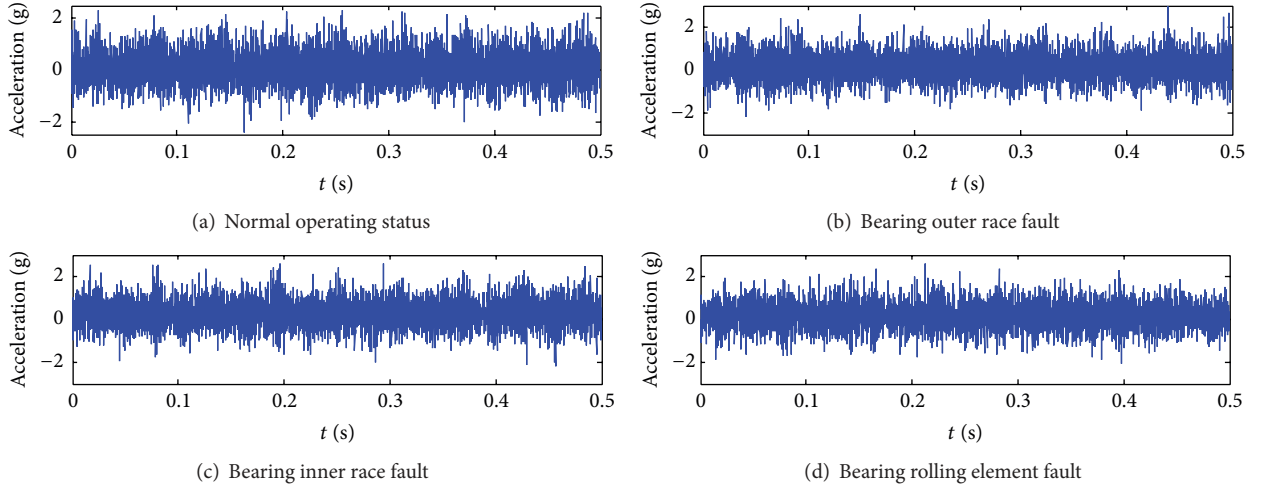


FIGURE 6: The original signal waveforms in different status of rolling bearing.

TABLE 2: The characteristic parameters in frequency domain.

Number	Feature expression
6	$p_1 = \frac{\sum_{k=1}^K s(k)}{K}$
7	$p_2 = \frac{\sum_{k=1}^K (s(k) - p_1)^2}{K - 1}$
8	$p_3 = \frac{\sum_{k=1}^K (s(k) - p_1)^3}{K (\sqrt{p_2})^3}$
9	$p_4 = \sqrt{\frac{\sum_{k=1}^K f_k^4 s(k)}{\sum_{k=1}^K f_k^2 s(k)}}$
10	$p_5 = \frac{\sum_{k=1}^K f_k s(k)}{\sum_{k=1}^K s(k)}$
11	$p_6 = \sqrt{\frac{\sum_{k=1}^K (f_k - p_5)^2 s(k)}{K}}$
12	$p_7 = \frac{\sum_{k=1}^K f_k^2 s(k)}{\sqrt{\sum_{k=1}^K s(k) \sum_{k=1}^K f_k^4 s(k)}}$

$s(k)$ is the frequency spectrum of signal $x(n)$, $k = 1, 2, 3, \dots, K$, K is the number of spectrum lines, and f_k is the frequency of the k th spectrum.

Figure 7. The decomposition process of vibration signals in other operating statuses is similar. The values of the statistical

TABLE 3: Experimental parameters of rolling bearing.

Experimental parameter	Value
Outer ring diameter	52 mm
Inner ring diameter	25 mm
Rolling ring diameter	7.5 mm
The number of rolling elements	12
Bearing contact angle	0°
Driving speed	1000 r/min
Bearing type	N205
Sampling frequency	10 KHz

features in 5 time domains and 7 frequency domains were extracted for each operating status. The first 8 IMF components containing the most fault information were used as the feature vector of the diagnostic model. SS-LLE was used for dimension reduction. A three-dimensional sensitive feature index chart after dimension reduction is shown in Figure 8.

5.2. *Experimental Results and Analysis.* The fault samples after dimension reduction were inputted into the IGA-MKSVM diagnosis model for classification recognition. The recognition results are shown in Figure 9. It is clear in Figure 9 that the average classification accuracy by using

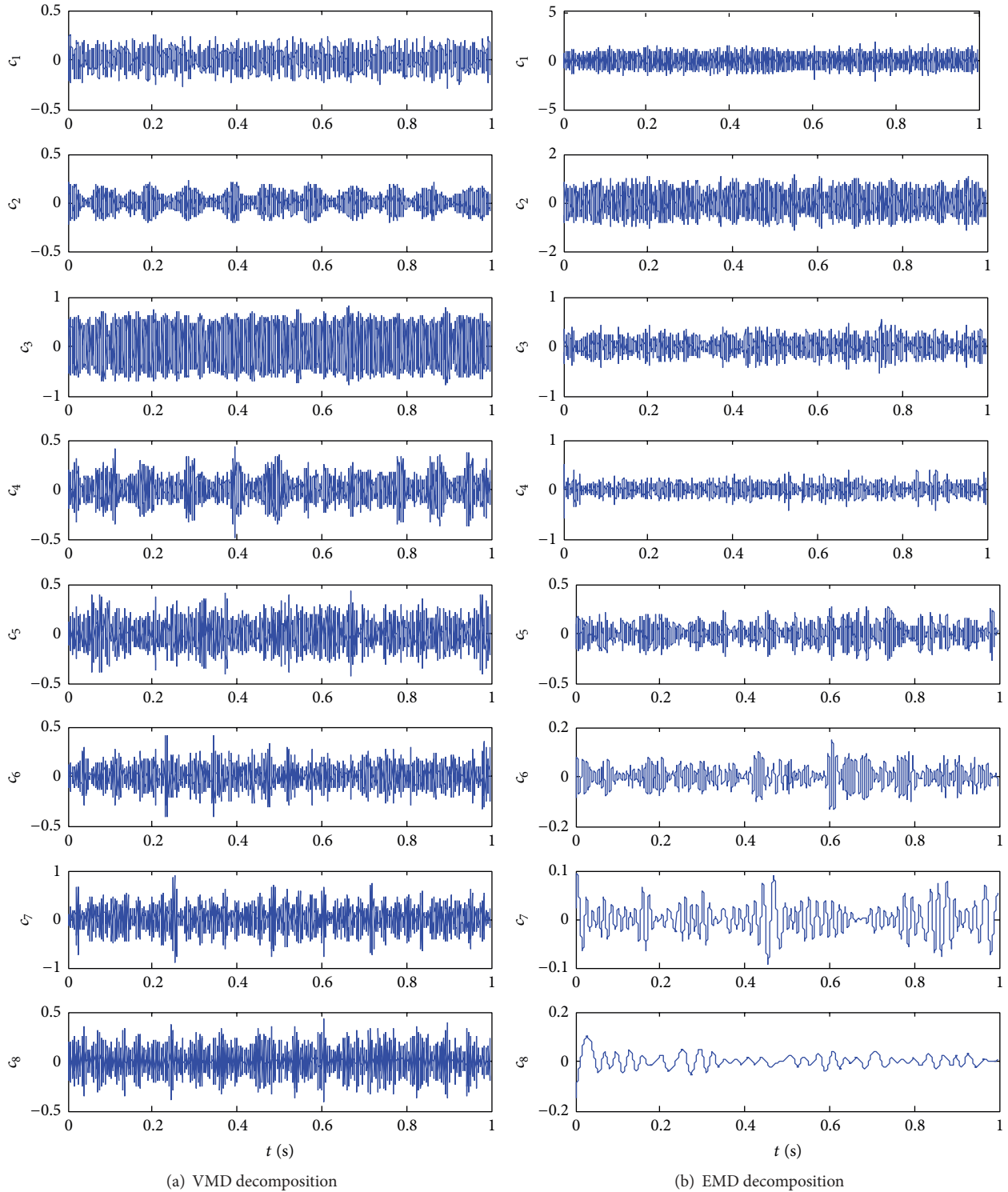


FIGURE 7: The decomposition results of VMD and EMD for a normal vibration signal.

VMD decomposition is 94.99%. It is higher than that of EMD decomposition which is 91.38%. Therefore, in the following experiments, VMD decomposition was used to construct the combined domain feature sets.

To further verify the advantage and stability of the IGA-MKSVM fault diagnosis model, the standard MKSVM, SVM with Gaussian kernel (GSVM), and SVM with polynomial kernel (PSVM) were chosen for comparison. We used fold

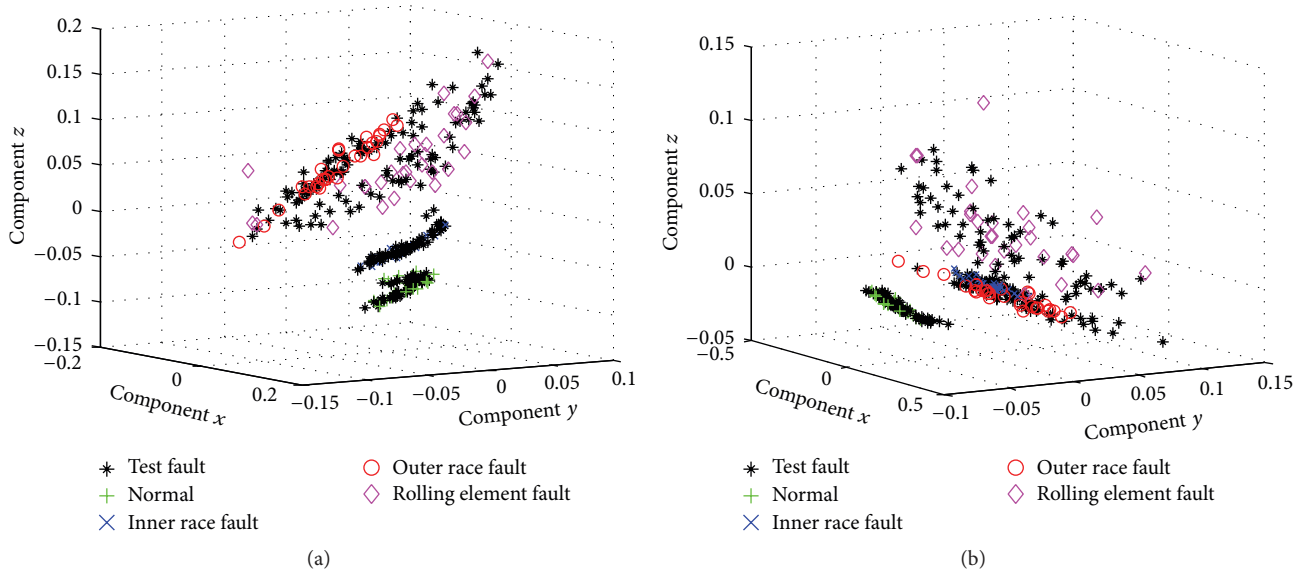


FIGURE 8: Comparison of dimension reduction after the decomposition of VMD and EMD.

TABLE 4: Comparison of classification accuracy with four different parameters.

Diagnostic model	Average percentage of correct recognition (%)			
	Normal	Outer ring fault	Inner ring fault	Rolling element fault
IGA-MSVM	100	96.67	93.33	90.00
MSVM	100	86.67	91.11	85.56
GSVM	100	81.11	84.44	74.44
PSVM	100	78.89	82.22	68.89

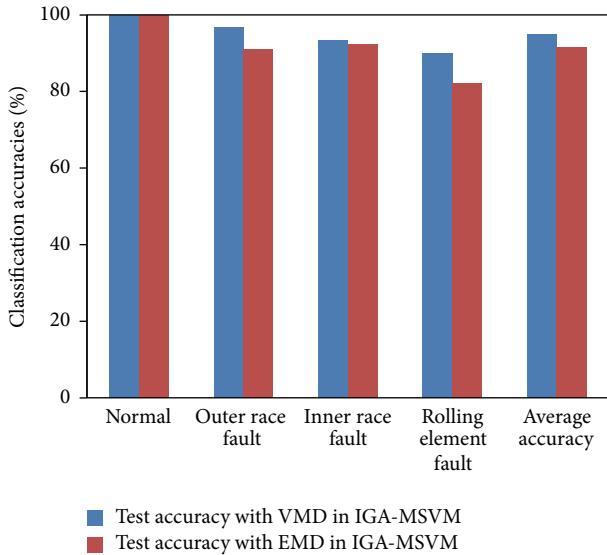


FIGURE 9: Average accuracy of classification by two different decomposition methods.

cross validation [23] with a total of 5 times of validation. The synthetic diagnostic results are shown in Table 4. It is clear in Table 4 that IGA-MKSVM diagnosis model achieves the

highest fault diagnosis accuracy. The recognition accuracy of the standard MKSVM is higher than that of SVM with a single kernel. Since the MKSVM combines the advantages of Gaussian kernel and polynomial kernel, it has higher recognition accuracy than SVM with a single kernel. Meanwhile, by obtaining the optimal control parameters through IGA, MKSVM model integrates the prior knowledge of fault features and improves the recognition performance.

To verify the robustness and generalization ability of IGA-MKSVM diagnostic model, stochastic noise signals with $\alpha = 0.1, 0.2, 0.3$ were added to the original signal. The original feature vector then becomes $X' = [x'_1, x'_2, \dots, x'_8]$, where $x'_k = [1 + \alpha \cdot \text{rand}(1)]x_k$, and x_k is the original time domain feature of IMF. The noisy feature vector was inserted into the IGA-MKSVM diagnostic model, standard MKSVM, GSVM, and PSVM. The results of them were compared by using 5-fold cross validation. The diagnostic results are shown in Table 5, which indicate that the IGA-MKSVM diagnostic model not only has relatively high diagnosis accuracy, but also has excellent noise immunity. When the extracted fault features are interfered at certain extent by noise, the diagnosis accuracy is still relatively high. In practical engineering applications, due to the interference of complicated environmental factors, it is hard to acquire sufficient fault information. Standard MKSVM and SVM with a single kernel are inadequate to engineering applications in such cases. However,

TABLE 5: Comparison of the noise immunity by using four different models for fault diagnosis.

Diagnostic model	Average correct ratio of recognition (%)		
	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$
IGA-MSVM	95.56	91.11	80.00
MSVM	88.89	83.33	71.11
GSVM	81.11	78.89	63.33
PSVM	77.78	68.89	48.89

IGA-MKSVM diagnostic model has noise immunity and good generalization ability, which is applicable to real fault diagnosis in engineering applications.

6. Conclusions

To achieve accurate and adaptive identification of mechanical faults, a new fault diagnosis method based on VMD and IGA optimized MSVM is proposed in this paper. IGA overcomes the immature convergence problem of traditional algorithms and is able to solve the initial parameter selection problem of MSVM. From this aspect, it makes MSVM more applicable, robust, and accurate. Experimental results demonstrate that, compared to traditional methods, the combination of VMD method and IGA-MSVM for fault diagnosis produces more repeatable results, with stronger generalization ability as well as anti-interference ability, providing a new effective method for mechanical fault diagnosis.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

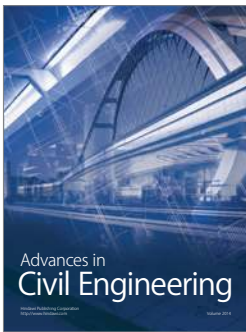
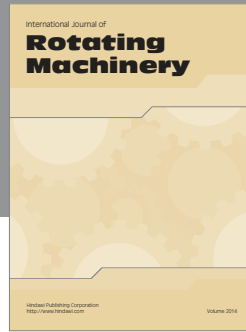
Acknowledgments

This research is supported by National Key Basic Research and Development Program (973 Program) (2015CB057702), the National Natural Science Foundation of China (Project no. 51205431), Chongqing Commission of Science and Technology Research Projects (Project no. KJ1401303), and the Research Foundation of Chongqing University of Science & Technology (Project no. CK2015Z19). Finally, the authors are very grateful to the anonymous reviewers for their helpful comments and constructive suggestions.

References

- [1] A. D. Nembhard, J. K. Sinha, A. J. Pinkerton, and K. Elbhah, "Combined vibration and thermal analysis for the condition monitoring of rotating machinery," *Structural Health Monitoring*, vol. 13, no. 3, pp. 281–295, 2014.
- [2] V. H. Nguyen, C. Rutten, and J.-C. Golinval, "Fault diagnosis in industrial systems based on blind source separation techniques using one single vibration sensor," *Shock and Vibration*, vol. 19, no. 5, pp. 795–801, 2012.
- [3] H. Sun, Z. He, Y. Zi et al., "Multiwavelet transform and its applications in mechanical fault diagnosis—a review," *Mechanical Systems and Signal Processing*, vol. 43, no. 1-2, pp. 1–24, 2014.
- [4] B. Muruganatham, M. A. Sanjith, B. Krishnakumar, and S. A. V. S. Murty, "Roller element bearing fault diagnosis using singular spectrum analysis," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 150–166, 2013.
- [5] J. Cheng, Y. Yang, and D. Yu, "The envelope order spectrum based on generalized demodulation time-frequency analysis and its application to gear fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 24, no. 2, pp. 508–521, 2010.
- [6] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [7] Q. Du and S. Yang, "Application of the EMD method in the vibration analysis of ball bearings," *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2634–2644, 2007.
- [8] S. J. Loutridis, "Damage detection in gear systems using empirical mode decomposition," *Engineering Structures*, vol. 26, no. 12, pp. 1833–1841, 2004.
- [9] J. Cheng, D. Yu, J. Tang, and Y. Yang, "Application of SVM and SVD technique based on EMD to the fault diagnosis of the rotating machinery," *Shock and Vibration*, vol. 16, no. 1, pp. 89–98, 2009.
- [10] P. Flandrin, G. Rilling, and P. Gonçalvés, "Empirical mode decomposition as a filter bank," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 112–114, 2004.
- [11] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.
- [12] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer Science & Business Media, 2000.
- [13] H. L. Ao, J. Cheng, K. Li, and T. K. Truong, "A roller bearing fault diagnosis method based on LCD energy entropy and ACROA-SVM," *Shock and Vibration*, vol. 2014, Article ID 825825, 12 pages, 2014.
- [14] A. Ben-Hur and W. S. Noble, "Kernel methods for predicting protein-protein interactions," *Bioinformatics*, vol. 21, supplement 1, pp. i38–i46, 2005.
- [15] D. Zheng, J. Wang, and Y. Zhao, "Non-flat function estimation with a multi-scale support vector regression," *Neurocomputing*, vol. 70, no. 1-3, pp. 420–429, 2006.
- [16] A. Rakotomamonjy, F. Bach, S. Canu, and Y. Grandvalet, "More efficiency in multiple kernel learning," in *Proceedings of the 24th International Conference on Machine Learning (ICML '07)*, pp. 775–782, ACM, Corvallis, Ore, USA, June 2007.
- [17] O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee, "Choosing multiple parameters for support vector machines," *Machine Learning*, vol. 46, no. 1-3, pp. 131–159, 2002.
- [18] W.-J. Lee, S. Verzakov, and R. P. W. Duin, "Kernel combination versus classifier combination," in *Multiple Classifier Systems*, vol. 4472 of *Lecture Notes in Computer Science*, pp. 22–31, Springer, Berlin, Germany, 2007.
- [19] J.-w. Luo and T. Wang, "Motif discovery using an immune genetic algorithm," *Journal of Theoretical Biology*, vol. 264, no. 2, pp. 319–325, 2010.
- [20] J. H. Holland, *Adaptation in Natural and Artificial Systems: an Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, The University of Michigan Press, 1975.

- [21] V. P. Sakthivel, R. Bhuvaneswari, and S. Subramanian, "Artificial immune system for parameter estimation of induction motor," *Expert Systems with Applications*, vol. 37, no. 8, pp. 6109–6115, 2010.
- [22] G. J. Tang and X. L. Wang, "Parameter optimized variational mode decomposition method with application to incipient fault diagnosis of rolling bearing," *Journal of Xi'an Jiaotong University*, vol. 49, no. 5, pp. 1–8, 2015.
- [23] J. Huang, X. Hu, and X. Geng, "An intelligent fault diagnosis method of high voltage circuit breaker based on improved EMD energy entropy and multi-class support vector machine," *Electric Power Systems Research*, vol. 81, no. 2, pp. 400–407, 2011.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

