Fault Diagnosis of Bearings using IG-SVM and EMD-SVD

Mohammad Heidari *
Department of Mechanical Engineering,
Aligudarz Branch, Islamic Azad University,
Aligudarz, Iran
E-mail: m.heidari@iau-aligudarz.ac.ir
*Corresponding author

Received: 13 July 2017, Revised: 23 August 2017, Accepted: 5 September 2017

Abstract: This paper present new method based on adaptive diagnosis, which can self-adaptively diagnose faults of bearings under varying operation conditions without any prior information. For this purpose, a new method using information-geometric support vector machine (IG-SVM) based on Empirical Mode Decomposition and Singular Value Decomposition (EMD–SVD) is presented. Firstly, the vibration signal is decomposed to singular features by the EMD-SVD. Then, the IG-SVM, which uses information geometry to modify SVM in a data-dependent way, is employed for fault clustering. The results show that the proposed method has an efficient approach for fault diagnosis of bearings.

Keywords: Bearing, Empirical mode decomposition, Fault diagnosis, Support vector machine

Reference: Mohammad Heidari, "Fault Diagnosis of Bearings Using IG-SVM and EMD-SVD", Int J of Advanced Design and Manufacturing Technology, Vol. 11/No. 1, 2018, pp. 71–80.

Biographical notes: M. Heidari received his PhD in Mechanical Engineering from Shahrekord University of Iran. He is currently Assistant Professor at the Department of Mechanical Engineering, Aligudarz Branch, Islamic Azad University, Aligudarz, Iran. His current research interest includes fault diagnosis of rotary machinery and artificial neural networks.

1 INTRODUCTION

By growing the industries, fault detection and diagnosis are as the main task of a monitoring system and have an effective role in the good operation and long life of systems [1]. The demand for safety and reliability in industries cases to an intelligent condition monitoring (CM) and fault diagnosis for the machineries [2]. CM plan gives an early alarm of component failures, then repairs or replacements can take place at the earliest time with the minimum loss of productivity [3]. Vibration analysis is a suitable method for various CM of rotating machineries [4]. Nowadays, a variety of data-driven methods exist for CM, such as PCA [5] and artificial neural network [6]. We know that, the operation conditions of machineries are usually unknown. Therefore, it is quite worthwhile to expand the diagnosis method for rotating machineries, which can self-adaptively diagnose faults under varying operation conditions without any prior cognition or human interference. Here, it is called "adaptive diagnosis". An adaptive diagnosis method can be provided with two characteristics:

1-Adaptively of varying operation conditions. Rotating machineries usually work under varying and severe operation conditions. The sensitivity of various features that are characteristics of machine health may vary considerably under different operation conditions [6]. Therefore, an adaptive diagnosis method ought to be self-adaptive under varying operation conditions.

2-Automaticity without any prior cognition or human interference.

If a fault diagnosis approach satisfies the above conditions, it is one sample of adaptive diagnosis methods. Utilization of the adaptive diagnosis method in CM systems can not only decrease the requirements of operators' experiences and skills, but also improve systems' efficiency and reduce the complexity and costs. This paper attempts to propose an adaptive diagnosis method for rotating machineries using vibration signal. We know that, feature extraction is important in the fault diagnosis. The Fourier transform (FT) has been used for extracting the features of stationary signals, as it could produce the statistical average of characteristics over the duration of the data.

Fourier transform is not suitable for nonlinear and non-stationary signals. In the most condition, vibration signals of rotating machineries are both nonlinear and non-stationary. In the most recent studies, time-frequency analysis methods have been used to study the vibration signal of the rotating machineries, such as bearings and gearboxes [7]. In the recent years, the wavelet transform has been used for signal processing. Wavelet transforms are capable of processing non-stationary signals, but they have some imperative deficiencies [8], [9]. Once the wavelet mother function is selected, it must be used to analyze all the vibration signal of rotary machinery [10]. Obviously, this non-adaptive nature cannot be suitable for adaptive diagnosis. To avoid these disadvantages of wavelet transforms, a self-adaptive method for nonlinear and non-stationary signals is needed.

Yang et al. [11] proposed the empirical mode decomposition (EMD), which uses local time and scale characteristics of the signal to decompose it into a number of intrinsic mode functions (IMFs). Recently, EMD has been used in fault diagnosis of the bearings [12]. Therefore, EMD can be used for feature extraction during the adaptive diagnosis. Meanwhile, considering that the vibration signal usually exhibits similar features to a periodic impulse when a local fault appears in the bearing, matrix singular value decomposition (SVD) techniques can be used to extract the features [13]. According to matrix theory, singular values are the intrinsic characteristics of a matrix with favorable stability, invariant ratio and rotation. Generally, SVD can be used as a tool for signal regularization, noise reduction, signal detection and estimation [14]. Usually, a form of phase space reconstruction from chaos theory is employed in SVD [15].

This requires reconstruction parameters, such as the delay time and embedding dimension, to be determined, which will involve extensive computations. Thus, this construction method is unavailable in a real-time application. However, EMD can decompose a signal into a number of IMFs, which can be used to construct the original matrix for SVD. Therefore, this paper proposes a hybrid feature extraction method that combines EMD and SVD to process nonlinear and non-stationary signals self-adaptively. After feature extraction, the next step is fault clustering. According to the characteristics of adaptive diagnosis, faults should be clustered automatically only based on the training data without manual intervention or any prior knowledge about operation conditions. In this case, the traditional discriminant analysis methods, such as Mahalanobis distance [16], FDA [17], which always require certain level of expertise and threshold settings, are not suitable. And yet, computation intelligence techniques are preferred in this situation [18].

Those computation intelligence techniques are data-driven and have been employed broadly in fault diagnosis, such as: Bayes net classifier [19], optimization algorithms [20], and artificial neural networks [21]. Among them, artificial neural networks (ANNs) are the most popular tools used by researchers. Unlike traditional methods which minimize the empirical training error, support vector machine (SVM), aims at minimizing an upper bound of the generalization error through maximizing the margin between the separating hyperplane and the data. What makes SVM attractive is the property of condensing information in the training data and providing a sparse representation by using a very small number of data points (SVs) [22]. The performance of SVMs largely depends on the kernel [23]. It is reported that choosing a kernel corresponds to a smoothness assumption of the discriminant function of the classifier. In case when there is some prior knowledge, it is available to choose a kernel [24]. However, in order to diagnose faults based on the training data without human interface or any prior knowledge, it is necessary to optimize the kernel in a data-dependent way.

In this work, an information-geometric method is employed. Amari [25] presented the Information geometry. It includes convex analysis and its duality as a special and important part. By elucidating its dualistic differential-geometrical structure, plenty of applications of information geometry have been widely prevailing in many fields [26]. When IG is selected in the area of SVM, in order to increase the margin or separability in the feature space without changing the volume of the entire space, it is efficient to enlarge volume elements locally in neighborhoods of support vectors which are located closely to the boundary surface. This makes it possible to enlarge the spatial resolution around the boundary so that the reparability between classes is increased.

For this reason, a conformal mapping of the input Riemannian space is selected. This will be realized approximately by a conformal transformation of a kernel. Therefore, based on the structure of the Riemannian geometry induced in the input space by the kernel, the SVM can be modified in a data-dependent way and the IG-SVM can be obtained [22]. In this paper, a new method using IG-SVM based on EMD-SVD is proposed for adaptive diagnosis of bearings. In this method, the EMD–SVD method is employed to self-adaptively extract features from the vibration signals. IG-SVM, which uses information geometry to modify SVM in a data-dependent way, is then utilized for fault clustering. This method adapts itself by the training data and can be used under varying operation conditions without any prior knowledge or human interface. The proposed method has shown the good performance in application on bearing fault diagnosis. Besides, current fault diagnosis methods always aim at certain objects. This paper is organized as follows: Section 2 introduces EMD, SVD, SVM, geometry of the SVM kernel, modifying SVM kernel using information geometry, as well as flow of the proposed adaptive diagnosis method; Section 3 describes the simulation experiment performed to test the IG-SVM; Section 4 describes the application cases of the proposed method for test rig of Shahrekord university; and Section 6 gives the conclusions of this paper.

2 METHODOLOGY

2.1 Empirical mode decomposition

The method of EMD is developed from the assumption that any signal consists of different simple intrinsic modes of oscillation. Each signal can be decomposed into a number of IMFs, each of which must satisfy the following definition [11]:

(1) In the whole data set, the number of extrema and the number of zero-crossings must either be equal or differ by at most one,

(2) At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. An IMF represents a simple oscillatory mode compared with the simple harmonic function. With this definition, any signal $x(t)$ can be decomposed as follows:

Step 1: Identify all the local extrema, and then connect all the local maxima by a cubic spline line to give the upper envelope.

Step 2: Repeat this procedure for the local minima to produce the lower envelope. Between them, the upper and lower envelopes should cover all the data.

Step 3: The mean of the upper and lower envelope values is designated as $m_1(t)$, the difference between the signal $x(t)$ and $m_1(t)$ is the first component $h_1(t)$:

$$x(t) - m_1(t) = h_1(t) \quad (1)$$

Ideally, if $h_1(t)$: is an IMF, then $h_1(t)$ is the first component of $x(t)$.

Step 4: if $h_1(t)$: is not an IMF, $h_1(t)$ is treated as the original signal and steps 1, 2 and 3 are repeated; then:

$$h_1(t) - m_{11}(t) = h_{11}(t) \quad (2)$$

After repeated sifting, i.e., up to k times, $h_{1k}(t)$ becomes an IMF, that is:

$$h_{1(k-1)}(t) - m_{1k}(t) = h_{1k}(t) \quad (3)$$

Then, it is designated as:

$$c_1(t) = h_{1k}(t) \quad (4)$$

The first IMF component from the original data $c_1(t)$ should contain the finest scale or the shortest period component of the signal.

Step 5: Separating $c_1(t)$ from $x(t)$:

$$r_1(t) = x(t) - c_1(t) \quad (5)$$

$r_1(t)$ is treated as the original data, and the above processes are repeated; therefore, the second IMF component $c_2(t)$ of $x(t)$ can be obtained. Repeating the process described above n times, n-IMFs of signal $x(t)$ can be obtained, Then:

$$\begin{cases} r_1(t) - c_2(t) = r_2(t) \\ \vdots \\ r_{n-1}(t) - c_n(t) = r_n(t) \end{cases} \quad (6)$$

The decomposition process can be stopped when $r_n(t)$ becomes a monotonic function from which no more IMFs can be extracted. Using this procedure, any signal can be decomposed [11]:

$$x(t) = \sum_{j=1}^{n} c_j(t) + r_n(t) \quad (7)$$

Thus, the signal is decomposed into n-empirical modes and a residue $r_n(t)$, which is the mean trend of $x(t)$. Each IMF $c_1(t), c_2(t), \dots c_n(t)$ contains lower-frequency oscillations than the prior-extracted one, while $r_n(t)$ represents the central tendency of signal $x(t)$. In this paper, the mirror periodic extending method (MPM) is employed to solve this problem.

2.2 SINGULAR VALUE DECOMPOSITION

SVD is a powerful and effective tool for feature extraction in linear algebra, and has been used for fault diagnosis in rotating machinery [15]. The SVD is defined as follows:
Let X denote a $m \times n$ matrix with $m \geq n$. According to the SVD theorem, the matrix X can be decomposed in the form [27]:

$$X = U\sigma V^T \quad (8)$$

Where $U(M \times M)$ and $V(N \times N)$ are orthogonal matrices and σ is an $M \times N$ diagonal matrix of singular values ($\sigma_{ij} \neq 0$ if $i = j$ and $\sigma_{11} \geq \sigma_{22} \geq \cdots \geq 0$). The columns of the orthogonal matrix U and V are called the left and right singular vectors, respectively. An important property of U and V is that they are mutually orthogonal. The singular values represent the importance of individual singular vectors in the composition of the matrix. In other words, a singular vector corresponding to the larger singular values have more information about the structure of the pattern embedded in the matrix than the other singular vectors.

2.3 CLASSIFICATION USING SVM

In this study, SVM is employed for fault classification. Suppose a pattern classifier, which uses a hyperplane to separate two classes of patterns based on given examples $D = \{(x_1, y_1), \dots, (x_i, y_i)\}$, where x is a vector in the input space $S = R^d$,$y \in \{-1, +1\}$ is class label and $i = 1, \dots, l$. A nonlinear SVM maps the input data x into a high dimensional feature space $F = R^N$ by using a nonlinear mapping $\emptyset(x)$. It then searches for a linear discriminant function:

$$f(x) = w.\emptyset(x) + b \quad (9)$$

In SVM, an optimum separating hyperplane that maximizes the margins is constructed by minimizing the following objective function:

$$Min\ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l} \xi_i \quad (10)$$

Subjected to:

$$y_i[w^T.\emptyset(x)_i + b] \geq 1 - \xi_i$$
$$\xi_i \geq 0, i = 1, \dots, l \quad (11)$$

Where ξ_i are slack variables to handle misclassifications, w is a weight vector, b, is scalar called bias and C is the cost parameter denoting the trade-off between the model complexity and the training error. In Eq. (11), $\emptyset(x)_i$ is a nonlinear function to map the input data to a high dimensional feature space in which the data can be separated linearly. To solve Eq. (11), one can take the Lagrangian, consider the necessary conditions for optimality and finally turn the minimization problem to the following dual form:

$$Max\ \sum_{i=1}^{l} \alpha_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} \alpha_i\alpha_j y_i y_j K(x_i, y_j)$$
$$\sum_{i=1}^{l} \alpha_i y_i = 0 \quad (12)$$

Subjected to:

$$0 \leq \alpha_i \leq C, i = 1, \dots, l \quad (13)$$

Where $K(x_i, x_j)$ is a kernel function representing inner product $\{\emptyset(x)_i, \emptyset(x)_i\}$ and α_i are Lagrangian multipliers. By solving Eq. (13), the optimal separating hyperplane is obtained as following:

$$b + \sum_{SV}\left[(\alpha_i y_i)K\left(x_i, x_j\right)\right] = 0 \quad (14)$$

And the optimal classifying rule is:

$$f = sgn(b + \sum_{SV}[(\alpha_i y_i)K\left(x_i, x_j\right)]) \quad (15)$$

Where SV denotes the support vectors for which the corresponding Lagrangian multipliers are positive. In this study, the following kernel functions are employed:
Polynomial:

$$K(x, \acute{x}) = (1 + x.\acute{x})^d \quad (16)$$

Gaussian radial basis function:

$$K(x,\acute{x}) = \exp(-\gamma\|x-\acute{x}\|^2) \quad (17)$$

Detailed explanation about the basic concepts of SVM theory can be found in Ref. [5]. For the employment of SVM, two parameters have to be selected in advance: C and γ. They could be optimized by n-fold cross-validation method [28].

2.4 GEOMETRY OF THE SVM KERNEL

To modify SVM kernel using information geometry, it is necessary to analyze the geometrical structure induced in the input space by a kernel as follows. The mapping $\emptyset(x)$ defines an embedding of S into F as a curved sub manifold. When F is a Euclidean or Hilbert space, a Riemannian metric is thereby induced in the space S, where the length of a small line element dx in S, is defined by the length in the larger space F. Denote by z the mapped pattern of x in the feature space, i.e., $z=\emptyset(x)$. A small vector dx is mapped to:

$$dz = \nabla\emptyset(x).dx = \sum_i \frac{\partial}{\partial x_i}\emptyset(x).dx_i \quad (18)$$

Where:

$$\nabla\emptyset(x) = \left(\frac{\partial}{\partial x_i}\emptyset(x)\right) \quad (19)$$

The squared length of $dz=(dz_\alpha)$ is written in the quadratic form as:

$$|dz| = \sum_\alpha (dz_\alpha)^2 = \sum_{i,j} g_{ij}(x)dx_i dx_j \quad (20)$$

Where:

$$g_{ij}(x) = \left(\frac{\partial}{\partial x_i}\emptyset(x)\right).\left(\frac{\partial}{\partial x_j}\emptyset(x)\right) \quad (21)$$

The dot denoting the summation over index α of ∅. The $n\times n$ positive-definite matrix $G(x)=(g_{ij}(x))$ is the Riemannian metric tensor induced in S. Here is a theorem:

$$g_{ij}(x) = \frac{\partial}{\partial x_i}\frac{\partial}{\partial \acute{x}_j}K(x,\acute{x})|_{\acute{x}=x} \quad (22)$$

From Eq. (12), we have:

$$\frac{\partial}{\partial x_i}\frac{\partial}{\partial \acute{x}_j}K(x,\acute{x}) = \nabla\emptyset(x).\nabla\emptyset(\acute{x}) \quad (23)$$

This proves the Eq. (22). As aforementioned, the Polynomial and Gaussian radial basis function are selected in this paper. Therefore, for Polynomial kernel, the induced Riemannian metric is:

$$g_{ij}(x) = d\delta_{ij} + x_i x_j d(d-1) \quad (24)$$

For Gaussian radial basis function, the induced Riemannian metric is:

$$g_{ij}(x) = 2\gamma.\delta_{ij} \quad (25)$$

The induced metric is Euclidean, which is translational and rotational invariant.

2.5 MODIFYING KERNEL USING IG

On the basis of geometry of the SVM kernel, to increase the margin or separability of classes, it is needed to enlarge the spatial resolution around the boundary surface in F. This idea can be implemented by a conformal transformation of kernel:

$$\widetilde{K}(x,\acute{x}) = c(x)c(\acute{x})K(x,\acute{x}) \quad (26)$$

With a properly positive scalar function $c(x)$. $\widetilde{K}(x,\acute{x})$ is called a conformal transformation of a kernel by factor $c(x)$. The nonlinear mapping $\emptyset(x)$ can be regarded as being modified to $\widetilde{\emptyset}(x)=c(x)\emptyset(x)$, which satisfies the Mercer positivity condition. The metric $\tilde{g}_{ij}(x)$ can be obtained:

$$\tilde{g}_{ij}(x) = c(x)^2 g_{ij}(x) + c_i(x)c_j(x) + 2c_i(x)c(x)K(x,x) \quad (27)$$

Where $c_i(x)=\frac{\partial c(x)}{\partial x_i}$ and $K_i(x,x) = \partial K(x,\acute{x}) / \partial x_i |_{\acute{x}=x}$. The last term is zero for the Gaussian RBF kernel. Therefore, if the function $c(x)$ is chosen in a way such that its value is large when x is close to the boundary and small otherwise, the idea of enlarging the spatial resolution around the boundary can be realized [23]. However, the boundary are unknown in practice, it is possible to use the empirical knowledge that the support vectors (SVs) are most likely located around the boundary, and choose $c(x)$ to have large values at SV positions. Taking the above analysis into consideration, $c(x)$ can be chosen as:

$$c(x) = \sum_{i\in SVs}\exp(-\|x-x_i\|^2/\tau_i^2) \quad (28)$$

Where the parameter τ is given by:

$$\tau_i^2 = \frac{1}{M}(\sum_\alpha \|x_\alpha - x_i\|^{2p})^{\frac{1}{p}} \quad (29)$$

The Eq. (29) runs over M SVs x_α that are nearest to x_i. When $p = 1$, τ_i^2 is the mean squared distance from x_i to its M nearest SV neighbors. The function $c(x)$ decreases exponentially with the distance to SVs. Out of the circles, the value of $c(x)$ is very small, so is its derivative [24]. In general, the SVM can be modified as follows:

1. Choose a kernel function K, train SVM and record the information of SVs,

2. Modify the kernel K according to the Eqs. (26), (27), (28), and obtain the modified kernel $\widetilde{K}$,

3. Train SVM with the modified kernel $\widetilde{K}$.

It is worth mentioning that repeating the above procedure would not improve the performance of SVM further [23]. In this way, the modified SVM can be self-adaptively based on the training data without any prior knowledge. Since the SVM is modified using information geometry, the modified SVM is called information-geometric SVM (IG-SVM) in this paper.

2.6 ADAPTIVE DIAGNOSIS METHOD

The proposed method employs self-adaptive algorithms in both feature extraction and fault clustering, which make the proposed method appropriate for the idea of "adaptive diagnosis". EMD self-adaptively decomposes the vibration signal into n-empirical modes (IMFs) and a residue r_n. Here, the MPM is employed to solve the end effects problem. The IMFs $c_1, c_2, \ldots c_n$ and r_n are constructed as a feature matrix. SVD is then selected to obtain an $n + 1$ dimensional feature vector of singular values. The initial SVM is trained with the primary kernel. Then, a modified kernel $\widetilde{K}$ by a conformal transformation is selected to construct the IG-SVM. IG-SVM can diagnose faults of bearings in a data-dependent way.

3.1 SIMULATION DATA

A simulated two-dimensional data set $z = (x, y)$ uniformly distributed in the region [-3, 3] × [-3, 3] where two classes are separated by a nonlinear boundary determined by $y = 1.52\sin(\frac{\pi}{2}x)$ was considered. The IG-SVM or SVM produces a new boundary to classify the data set. Misclassification happens when a pattern is in the region between the two boundaries. Therefore, success rate of classification can be used to measure the performance of the classifier. In this simulation, 200 training data were randomly and uniformly generated. Another 1800 testing data were also generated to test the performance of the IG-SVM or SVM.

3.2 SIMULATION RESULTS

In this paper two kernel functions such as RBF and polynomial were selected. In IG-SVM, p was generally assigned the value two, and M is chosen to be four. two hundred training data were used to train the SVM with a RBF kernel firstly. The parameters that must be determined are the kernel parameter γ and the cost C. A grid research method with cross-validation was employed to optimize the parameters. As a result, $C = 12$, and $\gamma = 24.42$. As shown in Table 1, the success rate of SVM was 92.14%, while the success rate of IG-SVM was 94.70%. The IG-SVM performed better than SVM. Besides, the number of SVs in SVM was 53 and in IG-SVM was 23.

This implies that the trained IG-SVM reduced the complexity of SVM. Meanwhile, with the purpose to verify whether the IG-SVM can perform well even if the kernel parameter is set badly, the classification results were also obtained when γ=10, 20, 30, 40, as shown in Table 1 and Figure 1. In general, when the value of γ is further away from the optimal value, the SVM performed worse. However, the IG-SVM always performed well. when γ=10, the success rate of SVM even dropped down to less than 91%, the IG-SVM still classified the testing data successfully at a high level at 91.77%.

Table 1 Rate of classification with SVM and IG-SVM

	SVM-RBF		IG-SVM-RBF	
	Test success	Percentage	Test success	Percentage
γ=10	1819	90.33%	1866	91.77%
γ=20	1822	88.15%	1891	93.13%
γ=24.42	1888	92.14%	1853	94.70%
γ=30	1815	90.99%	1890	93.08%
γ=40	1832	92.90%	1887	94.01%

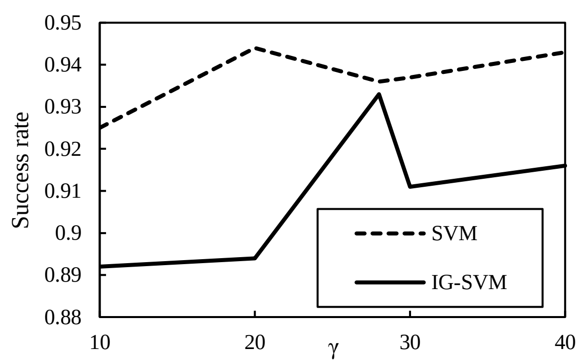

Fig. 1 Comparison of success rates obtained by SVM and IG-SVM with an RBF kernel

Similar to the RBF kernel, the simulation results of SVM and IG-SVM with a Polynomial kernel were obtained when d=2, 3, 4, 5 as shown in Figure 2 and Table 2. All success rates of SVM were below 86%. Compared with the results of the RBF kernel, it is obviously that the Polynomial kernel was not very suitable for the simulation data. However, even with a bad kernel, the IG-SVM improved the performance of SVM largely. When d=4, the success rate of IG-SVM reached to 92.63%. It can be observed that compared with the SVM, the IG-SVM has certain robustness to the inappropriate kernels.

Table 2 Rate of classification with SVM and IG-SVM with a polynomial kernel

	SVM		IG-SVM	
	Test success	Percentage	Test success	Percentage
d=2	1880	84.80%	1896	88.42%
d=3	1878	84.14%	1854	91.03%
d=4	1899	86.20%	1866	91.60%
d=5	1809	85.84%	1835	90.32%

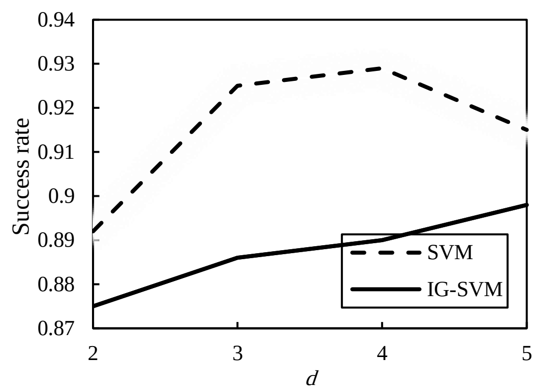

Fig. 2 Comparison of success rates obtained by SVM and IG-SVM with a polynomial kernel

4.1 BEARING FAULT DIAGNOSIS

To verify that the propose method can diagnose faults under varying operation conditions, vibration signals of bearings were used. The datasets were acquired from a test rig in the Case Western Reserve University Bearing Data Center. In this experiment, as Table 3 shows, it has been tested 6205-2RS JEM SKF deep-groove ball bearings with motor speeds of 1796, 1772, 1748 and 1722 r/min. Under varying operation conditions, three common types of bearing faults were tested: inner race fault, outer race fault, and rolling element fault. The vibration signals were acquired with a sampling rate of 12 kHz. Datasets under 4 different operation conditions were investigated to verify the proposed method. Under each operation condition, there are 50 training samples: 20 from normal case, 10 from inner race fault, 10 from outer race fault, and 10 from rolling element fault; as well as 100 test samples: 40 from normal case, 20 from inner race fault, 20 from outer race fault, and 20 from rolling element fault. The trained IG-SVM (or SVM) under each operation condition would be employed to diagnose faults under all operation conditions. In this way, it can be verified that the proposed method can be used under varying operation conditions self-adaptively without any prior knowledge or human interface.

Table 3 The case Western Reserve University bearing data

Type	6205-2RS JEM SKF, deep groove ball bearing							
Fault type	Motor speed (r/min)							
	1796		1772		1748		1722	
	T. S	Te. S	T. S	Te. S	T. S	Te. S	T. S	Te. S
N	20	40	20	40	20	40	20	40
IRF	10	20	10	20	10	20	10	20
ORF	10	20	10	20	10	20	10	20
BF	10	20	10	20	10	20	10	20
Total	50	100	50	100	50	100	50	100

Firstly, the vibration signal of each dataset was decomposed into n-empirical modes and a residue $r_n(t)$. Information concerning the health condition of the bearings is contained in the high frequency bands. Generally, n was always assigned the value 5. If the number of IMFs is less than 5, it can be made up the difference with zero vectors. Secondly, the constructed feature matrix was decomposed to obtain a 6-dimensional feature vector of singular values. Table 4 shows the results of fault diagnosis using the IG-SVM and SVM.

Table 4 Rate of bearing fault diagnosis

Trained model under motor speed (r/min)		No. of SVs	Test samples under motor speed (r/min)			
			1796	1772	1748	1722
1796	SVM	46	100%	100%	99%	99%
	IG-SVM	26	100%	100%	100%	100%
1772	SVM	47	94%	100%	100%	89%
	IG-SVM	27	96%	100%	100%	95%
1748	SVM	49	93%	98%	100%	91%
	IG-SVM	21	97%	100%	100%	97%
1722	SVM	47	96%	98%	93%	100%
	IG-SVM	20	97%	100%	98%	100%

Generally, if the trained model and test samples were under the same operation condition, the success rates were all at a level of 100%. This owns to the hybrid feature extraction method (EMD-SVD) which can extract the remarkable features that reflect the health state of bearing. However, if the trained model and test samples were under different operation conditions, misclassification emerged. For example, when the trained SVM under the motor speed of 1772 r/min was employed for fault diagnosis under the motor speed of 1722 r/min, the success rate was only 89%. This rate is unacceptable in real-world applications.

Traditionally, to deal with this issue, it is needed to intervene on the diagnosis model with prior knowledge about the current operation condition. However, without this need, the IG-SVM performed well in this situation. Although the operation condition of test samples was unknown, the IG-SVM had adapted itself in a data-dependent way. The success rates of fault diagnosis using the IG-SVM were no less than 94%. Besides, the trained IG-SVMs reduced the number of SVs too, which also indicates that the trained IG-SVM reduced the complexity of SVM. In summary, the proposed method using IG-SVM based on EMD-SVD can adaptively diagnose bearing faults under varying operation conditions without any prior knowledge or human interface.

4.2 SHAHREKORD UNIVERSITY TEST RIG FOR FAULT DIAGNOSIS

The efficiency of the proposed method has been verified in application on bearing fault diagnosis. Nevertheless, since fault diagnosis of many rotating machineries mainly possess vibration signals, it is possible to extend the proposed method to more objects. Furthermore, this paper also tried to apply it on Shahrekord University test rig. This test rig, as shown in Fig. 3, was tested and analysed to verify the proposed method. However, owing to some unsolvable restrictions, the proposed method can only be validated with the vibration data under a stable operation condition.

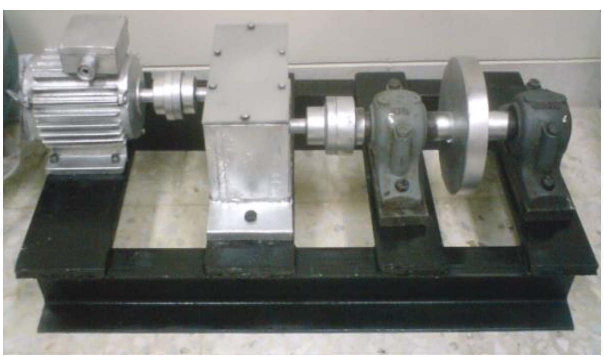

Fig. 3 Fault simulator set up in Shahrekord University

The experimental setup at Shahrekord University to collect dataset consists of a one-stage gearbox with spur gears, a flywheel and an electrical motor. Electrical motor, gearbox and flywheel are attached together through flexible couplings as shown in Fig. 3. 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 second. 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 were 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. Here, 150 training samples were used: 30 from normal case (N), 30 from inner race fault (IRF), 30 from outer race fault (ORF), 30 from ball fault (BF) and 30 from combined fault (CF) in bearing. Similar to the application on bearing fault diagnosis, the results of the proposed method could be obtained, as shown in Table 5. The results show that when the trained SVM was employed, there is a high level of false alarm rate (13/16); and yet, the IG-SVM could reduce the false alarms and diagnose accurately. It implies that the application on Shahrekord test rig fault diagnosis with the proposed method is feasible and efficient.

Table 5 Results of Shahrekord test rig

		SVM	**IG-SVM**
Success rate of classification	N	13/30	16/30
	IRF	16/30	16/30
	ORF	14/30	15/30
	BF	14/30	25/30
	CF	18/30	27/30
	Total	43/48	47/48
	Percentage	89.58%	97.92 %
NO. of SVs		28	16

5 CONCLUSIONS

Fault diagnosis is the main technique for condition monitoring systems. Bearings usually work under varying and severe operation conditions. This paper attempts to propose a conception: adaptive diagnosis, which means the method can diagnose faults automatically under varying operation conditions without any prior knowledge or human interface. This paper presents a new method using IG-SVM based on EMD-SVD for fault diagnosis of bearings. This method employs self-adaptive algorithms in both feature extraction and fault clustering, and can adapt itself automatically during fault diagnosis under varying operation conditions. During feature extraction, it extracts those features that evidently reflect the health state of the bearings, because EMD is self-adaptive in decomposing nonlinear and non-stationary signals and SVD extracts the intrinsic characteristics of a matrix with favourable stability. Also, during fault clustering, the IG-SVM can not only cluster faults automatically as same as SVM, but also improve the performance of SVM in a data-dependent way.

The feasibility and efficiency of IG-SVM was validated by a simulation experiment. The results show that the IG-SVM can not only perform better than SVM, but also have certain robustness to selection of kernel function and parameter configuration compared with SVM. Those advantages will facilitate the widespread application of IG-SVM. Then, the proposed adaptive diagnosis method was applied on bearing fault diagnosis under varying operation conditions and performed effectively. Additionally, to extend the application of this method, this paper also took a try on Shahrekord test rig fault diagnosis.

REFERENCES

[1] Wang, J., Gao, R. X., and Yan, R., “A Hybrid Approach to Bearing Defect Diagnosis in Rotary Machines”, CIRP Journal of Manufacturing Science and Technology, Vol. 5, No. 4, 2012, pp. 357–365.

[2] Lou, X., Loparo, K. A., “Bearing Fault Diagnosis Based on Wavelet Transform and Fuzzy Inference”, Mechanical Systems and Signal Processing, Vol. 18, No. 5, 2004, pp. 1077–1095.

[3] Konar, P., Chattopadhyay, P., “Bearing Fault Detection of Induction Motor Using Wavelet and Support Vector Machines (SVMs)”, Applied Soft Computing, vol. 11, No. 6, 2011, 2011, pp. 4203–4211.

[4] Bayram, D., Şeker, S., “Wavelet Based Neuro-Detector for Low Frequencies of Vibration Signals in Electric Motors”, Applied Soft Computing, Vol. 13, No. 5, 2013, pp. 2683–2691.

[5] Li, Z., He, Z., Zi, Y., and Jiang, H., “Rotating Machinery Fault Diagnosis Using Signal-adapted Lifting Scheme”, Mechanical Systems and Signal Processing, Vol. 22, No. 3, 2008, pp. 542–556.

[6] Chattopadhyay, P., “Bearing Fault Detection of Induction Motor Using Wavelet and Support Vector Machines (SVMs)”, Applied Soft Computing, Vol. 11, No. 6, 2011, pp. 4203–4211.

[7] Sanz, J., Perera, R., and Huerta, C., “Fault Diagnosis of Rotating Machinery Based on Auto-Associative Neural Networks and Wavelet Transforms”, Journal of Sound and Vibration, Vol. 302, No. 4-5, 2007, pp. 981–999.

[8] Kankar, P. K., Satish, C. S., and Harsha, S. P., “Fault Diagnosis of Ball Bearings Using Continuous Wavelet Transform”, Applied Soft Computing, Vol. 11, No. 2, 2011, pp. 2300–2312.

[9] Prabhakar, S., Mohanty, A. R., and Sekhar, A. S., “Application of Discrete Wavelet Transform for Detection of Ball Bearing Race Faults”, Tribology International, Vol. 35, No. 12, 2002, pp. 793– 800.

[10] Zamanian, A. H., Ohadi, A., “Gear Fault Diagnosis Based on Gaussian Correlation of Vibrations Signals and Wavelet Coefficients”, Applied Soft Computing, Vol. 11, No. 8, 2011, pp. 4807–4819.

[11] Yang, Y., He, Y., Cheng, J., and Yu, D., “A Gear Fault Diagnosis Using Hilbert Spectrum Based on MODWPT and a Comparison with EMD Approach”, Measurement, Vol. 42, No. 4, pp. 542–551, 2009.

[12] Rafiee, J., Rafiee, M. A., and Tse, P. W., “Application of Mother Wavelet Functions for Automatic Gear and Bearing Fault Diagnosis”, Expert Systems with Applications, Vol. 37, No. 6, 2010, pp. 4568–4579.

[13] Hajnayeb, A., Ghasemloonia, A., Khadem, S. E., and Moradi, M. H., “Application and Comparison of an ANN-Based Feature Selection Method and the Genetic Algorithm in Gearbox Fault Diagnosis”, Expert Systems with Applications, Vol. 38, No. 8, 2011, pp. 10205–10209.

[14] Rafiee, J., Arvani, F., Harifi, A., and Sadeghi, M. H., “Intelligent Condition Monitoring of a Gearbox Using Artificial Neural Network”, Mechanical Systems and Signal Processing, Vol. 21, No. 4, 2007, pp. 1746–1754.

[15] Abbasion, S., Rafsanjani, A., Farshidianfar, A., and Irani, N., “Rolling Element Bearings Multi-Fault Classification Based on the Wavelet Denoising and Support Vector Machine”, Mechanical Systems and Signal Processing, Vol. 21, No. 7, 2007, pp. 2933–2945.

[16] Chen, D., Wang, W. J., “Classification of Wavelet Map Patterns Using Multi-Layer Neural Networks for Gear Fault Detection”, Mechanical Systems and Signal Processing, Vol. 16, No. 4, 2002, pp. 695–704.

[17] Jiang, Y., Tang, B., Qin, Y., and Liu, W., “Feature Extraction Method of Wind Turbine Based on Adaptive Morlet Wavelet and SVD”, Renewable Energy, Vol. 36, pp. 2011, 2146-2153.

[18] Cherkassky, V., Ma, Y., “Practical Selection of SVM Parameters and Noise Estimation for SVM Regression”, Neural Networks, Vol. 17, No. 1, 2004, pp. 113-126.

[19] Wang, Y., He, Z., and Zi, Y., “Enhancement of Signal Denoising and Multiple Fault Signatures Detecting in Rotating Machinery Using Dual-Tree Complex Wavelet Transform”, Mechanical Systems and Signal Processing, Vol. 24, 2010, pp. 119–137.

[20] Nikolaou, N. G., Antoniadis, I. A., “Rolling Element Bearing Fault Diagnosis Using Wavelet Packets”, NDT&E International, Vol. 35, No. 3, 2002, pp. 197-205.

[21] Barakat, M., Druaux, F., Lefebvre, D., Khalil, M., and Mustapha, O., “Self Adaptive Growing Neural Network Classifier for Faults Detection and Diagnosis”, Neurocomputing, Vol. 74, No. 18, 2011, pp. 3865–3876.

[22] Li, N., Zhou, R., and Zhao, X. Z., "Mechanical Faulty Signal Denoising Using a Redundant Non-Linear Second-Generation Wavelet Transform", Proceedings of the Institution of Mechanical Engineer. Part C: Journal of Mechanical Engineering Science, Vol. 225, 2011, pp. 799–808.

[23] Xian, G. M., Zeng, B. Q., "an Intelligent Fault Diagnosis Method Based on Wavelet Packet Analysis and Hybrid Support Vector Machines", Expert Systems with Applications, Vol. 36, No. 10, 2009, pp. 2131–12136.

[24] Kumar, R., Singh, M., "Outer Race Defect Width Measurement in Taper Roller Bearing Using Discrete Wavelet Transform of Vibration Signal", Measurement, Vol. 46, No. 1, 2013, pp. 537- 545.

[25] Paya, B. A., Esat, I. I., and Badi, M. N. M., "Artificial Neural Network Based Fault Diagnostics of Rotating Machinery Using Wavelet Transforms as a Preprocessor", Mechanical Systems and Signal Processing, 1997, Vol. 11, No. 5, 1997, pp. 751-765.

[26] Chen, P. H., Lin, C. J., and Scholkopf, B., "A Tutorial on V-Support Vector Machines", Applied Stochastic Models in Business and Industry, Vol. 21, No. 2, 2005, pp. 111-136.

[27] Smola, A. J., Scholkopf, B., and Muller, K. R., "The Connection between Regularization Operators and Support Vector Kernels", Neural Networks, Vol. 11, No. 4, 1998, pp.637–649.

[28] Keskes, H., Braham, A., and Lachiri, Z., "Broken Rotor Bar Diagnosis in Induction Machines Through Stationary Wavelet Packet Transform and Multiclass Wavelet SVM", Electric Power Systems Research, Vol. 97, 2013, pp. 151–157.