2014

Fault Diagnosis for Drivetrain Gearboxes Using PSO-Optimized Multiclass SVM Classifier

Dingguo Lu
University of Nebraska-Lincoln, stan1860@huskers.unl.edu

Wei Qiao
University of Nebraska-Lincoln, wqiao@engr.unl.edu

PES General Meeting | Conference & Exposition, 2014 IEEE
Year: 2014
Pages: 1 - 5, DOI: 10.1109/PESGM.2014.6939892

Fault Diagnosis for Drivetrain Gearboxes Using PSO-Optimized Multiclass SVM Classifier

Dingguo Lu, *Student Member, IEEE*, and Wei Qiao, *Senior Member, IEEE*

***Abstract*—A novel method consisting of an adaptive feature extraction scheme and a particle swarm optimization (PSO)-optimized multiclass support vector machine (SVM) classifier is proposed for condition monitoring and fault diagnosis of drivetrain gearboxes in variable-speed operational conditions. The adaptive feature extraction scheme consists of an adaptive signal resampling algorithm, a frequency tracker, and a feature generation algorithm for effective extraction of the features of gearbox faults from the stator current signal of the AC electric machine connected to the gearbox. The multiclass SVM classifier is designed to identify different faults in the gearbox according to the fault features extracted. The PSO algorithm is utilized to optimize the parameter setting of the SVM classifier to obtain the best classification accuracy. The proposed method is testified on a drivetrain gearbox connected with a permanent-magnet synchronous machine with three different faults. Experimental results show that the faults can be effectively classified by the proposed method.**

***Index Terms*—Condition monitoring, drivetrain, fault diagnosis, gearbox, multiclass classification, particle swarm optimization (PSO), support vector machine (SVM).**

I. Introduction

GEARBOXES are widely used in the drivetrains of various mechanical and electromechanical systems, such as wind energy systems, (hybrid) electric vehicles, electric train, aircrafts, etc. It is beneficial to develop reliable condition monitoring and fault diagnosis (CMFD) technologies for the safe operation of these systems [1]-[4].

Prior work has demonstrated the effectiveness of current-based methods for CMFD of gearboxes [5]-[7]. However, the previous work assumed that the gearbox had been in faulty condition but did not address the diagnostic task when the condition of the gearbox is unknown. As one of the pattern recognition methods, the support vector machine (SVM) offers an effective means to solve the problem of fault identification [8]. The SVM possesses useful properties for the problems of classification in terms of the complexity and efficiency of computation, uniqueness of solution, and simplicity of implementation. Owing to these merits, a SVM-based classifier [9] was proposed for gearbox fault diagnosis.

The drivetrain gearboxes in industrial systems suffer a variety of faults. Hence, the fault classification of drivetrain gearboxes is a multiclass problem. The SVM was originally designed for binary classification [10]. Some effort has been made to extend it for multiclass classification [11]-[13]. Although the SVM algorithm is effective to perform pattern recognition in many applications, such as facial expression analysis, texture classification, and image clustering, the optimal values of the SVM parameters are hard to determine, which limits the use of the SVMs. Generally, the parameter setting of an SVM significantly influences the classification accuracy of the SVM. The key parameters of an SVM include the kernel parameters and penalty parameter. The existing approaches to determining these parameters are based on prior knowledge, user expertise, or experimental trial. However, there is no general consensus for setting the parameters of an SVM and many opinions for choosing the optimal SVM parameters are contradictory [14]. Furthermore, the SVM generalization performance depends on all of the parameters simultaneously. This makes the optimal parameter selection even more complicated. A separate optimization for each parameter does not necessarily ensure an optimized SVM model [15]. Grid search [16] is a conventional method that has been used for finding the optimal parameters of an SVM. However, this method is considered expensive in terms of computational costs and data requirements. Hence, a computationally effective method that is capable of searching for the optimal values of all the parameters simultaneously is desired for the optimal parameter setting of multiclass SVM classification.

This paper proposes a novel method of using the particle swarm optimization (PSO) algorithm to optimize a multiclass SVM classifier for fault classification of drivetrain gearboxes. An adaptive feature extraction algorithm is proposed to extract the fault-related features from nonstationary current signals. A multiclass SVM classifier is designed to classify various gearbox faults. The PSO algorithm is utilized to optimize the parameter setting of the SVM classifier. Experimental studies are carried out to validate the proposed method for a drivetrain gearbox.

II. Characteristic Frequencies of Drivetrain Gearboxes

If a drivetrain gearbox is connected to an AC electric machine, the faults in the gearbox will change the sideband distribution in the frequency spectrum of the current signals of the AC machine [17]. Mathematical expression has been derived to show how the torsional vibrations related to the

This work was supported in part by the U.S. National Science Foundation under Grant ECCS-1308045 and the U.S. Department of Energy under Grant DE-EE0001366.

D. Lu and W. Qiao are with the Power and Energy Systems Laboratory, Department of Electrical Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588-0511 USA (e-mail: stan1860@huskers.unl.edu, wqiao3@unl.edu).

gearbox faults affect the current signals and demonstrated the characteristic frequencies of gearbox faults in the current frequency spectrum [5], [18].

This paper considers a drivetrain consisting of a two-stage gearbox connected with a permanent-magnet synchronous machine (PMSM), as illustrated in Fig. 1, where z_1–z_4 denote the tooth numbers of the four gears in the gearbox. The characteristic frequencies of the gearbox vibration include the input shaft frequency f_1, pinion shaft frequency f_2, output shaft frequency f_3, and two gear meshing frequencies f_{m1} and f_{m2}. These characteristic frequencies then modulate the stator currents of the PMSM and generate sidebands across the dominant components of the currents [5]:

$$f_{sideband} = k f_s \pm l f_1 \pm m f_2 \pm n f_3 \quad (1)$$

where f_s is the fundamental frequency of the stator current, and $k, l, m, n = 1, 2, 3, \cdots$.

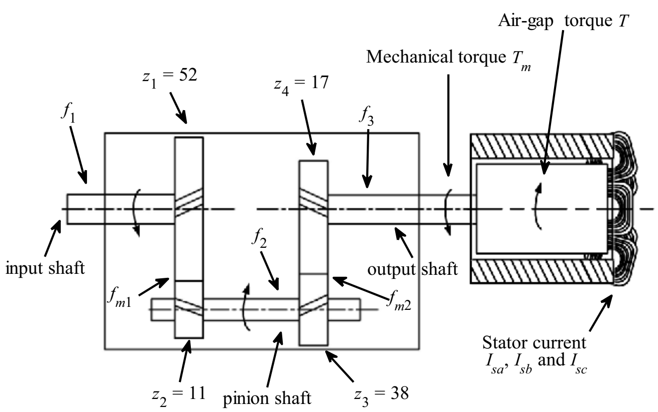

Fig. 1. Schematic of a drivetrain consisting of a two-stage gearbox connected to a PMSM with characteristic vibration frequencies.

III. Adaptive Feature Extraction

A. Adaptive Signal Resampling

The characteristic frequencies of gearbox faults in the stator current are related to gearbox shaft speeds, and become nonstationary when the shaft speeds vary with time. This paper uses an adaptive signal resampling algorithm to convert the current samples from a fixed sampling rate to an adaptive sampling rate to make the values of the objective frequencies, such as the fault characteristic frequencies, constant, as illustrated in Fig. 2. The details of the algorithm are described in [5]. By using the adaptive resampling algorithm, the time-varying characteristic frequencies of the gearbox in the stator current are converted to constant values in the PSD spectrum of the resampled signal. The magnitudes of certain frequency components in the resultant PSD spectrum can be then used to generate features to evaluate the condition of the gearbox.

B. Frequency Tracker

The characteristic frequencies of the gearbox in Fig. 1 are related to each other because of their mechanical connection, and the rotating frequency of gearbox output shaft is identical to that of the PMSM shaft.

$$f_2 = \frac{z_4}{z_3} f_3 \quad (2)$$

$$f_1 = \frac{z_2}{z_1} f_2 = \frac{z_2}{z_1}\frac{z_4}{z_3} f_3 \quad (3)$$

The objective features of faults are different among different data records, as the operating points, e.g., the speed and load, may vary over time. To facilitate the fault feature extraction in this work, a frequency tracker is developed to solve this problem by utilizing the mechanical relationship among the gearbox shafts. In the proposed frequency tracker, a frequency detector is designed to detect the fundamental frequency of the resampled stator current signal. Then, according to (2) and (3), the characteristic frequencies, f_1, f_2 and f_3, of the gearbox can be calculated. After that, an objective frequency extractor is designed to extract the objective frequency components with their magnitudes according to (1). In this works, the objective frequency components are the sidebands around the fundamental frequency of the resampled stator current.

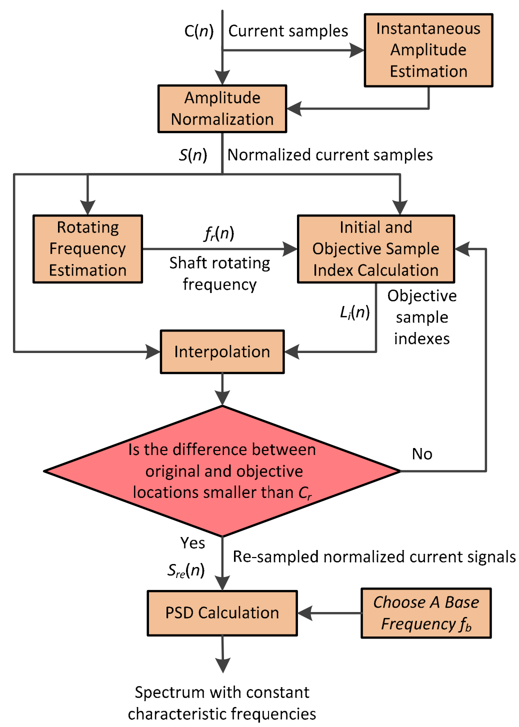

Fig. 2. Schematic of the adaptive signal resampling algorithm.

C. Frequency Generation

This paper uses the magnitudes of two groups of sidebands around the fundamental frequency calculated from the frequency tracker to generate the features related to faults. The first group is the sidebands caused by the first to forth order of the gearbox input shaft rotating frequency, which are expressed as $f_s \pm l f_1$, where $l = 1, 2, 3$ and 4. The second group is the sidebands caused by the three gearbox shaft rotating frequencies, which are expressed as $f_s \pm f_1$, $f_s \pm f_2$ and $f_s \pm f_3$. The magnitudes of these two groups of sidebands are first normalized with respect to the magnitude of the fundamental frequency component. Two features are then generated for each group from the normalized sideband magnitudes: the standard deviation defined in (5) and the summation defined in (6).

$$\sigma = \left[\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2 \right]^{\frac{1}{2}} \tag{5}$$

$$M = \sum_{i=1}^{N} x_i \tag{6}$$

where x_i $(i = 1, 2, \dots, n)$ is the normalized magnitude of each sideband, $\bar{x}$ is the mean value of x_i, and N is number of sideband pairs being processed.

The adaptive signal resampling algorithm, the frequency tracker, and the feature generation algorithm constitute the proposed adaptive feature extraction scheme in Fig. 3.

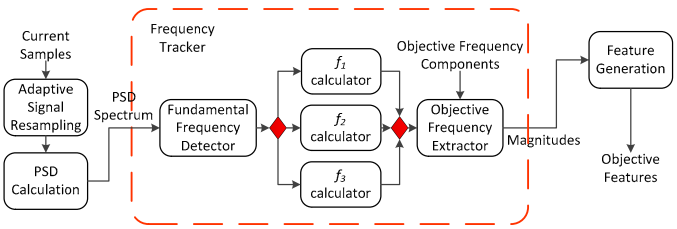

Fig. 3. Schematic diagram of the proposed feature extraction scheme.

IV. Multiclass SVM Classification

A. Principles of SVM

The basic idea of the SVM-based classification for a binary problem is to construct a hyper-plane as the decision plane, which separates the positive (+1) and negative (-1) classes with the largest margin (Fig. 4). The margin is the sum of the distances from the hyper-plane to the two boundaries constructed by the closest data points of the two classes. These closest data points are defined as the Support Vectors.

Suppose that there is a given training data set $G = \{(x_i, y_i), i = 1 \dots P\}$. Each sample $x_i \epsilon R^D$ belongs to a class $y_i \in \{+1, -1\}$. The hyper-plane of the SVM can be expressed as follows:

$$\boldsymbol{\omega} \cdot x + \mathbf{b} = 0 \tag{7}$$

where ω is a weight vector and b is a bias vector.

Thus, the following decision function can be used to classify any data point x in either class "+1"or "-1":

$$f(x) = \operatorname{sgn}(\boldsymbol{\omega} \cdot x + \mathbf{b}) \tag{8}$$

where $\operatorname{sgn}(\cdot)$ is the operation to find the sign of a value.

According to (8), a SVM-based classifier is as follows:

$$f(x) = \operatorname{sgn}\left[\sum_{i=1}^{P} a_i y_i K(x, x_i) + \mathbf{b} \right] \tag{9}$$

which is subject to

$$\sum_{i=1}^{l} \alpha_i y_i = 0 \tag{10}$$

where $\alpha_i \geq 0$ is a Lagrange multiplier, $K(x, x_i)$ is the kernel function.

The kernel function used in the SVM application is to map the input vectors into a higher-dimensional feature space through some nonlinear separating hyper-plane and, thus, makes the data linearly separable in the feature space although the original input vectors are nonlinearly separable in the input space [8]. Among all the available kernel functions for SVMs, the radial basis function (RBF) kernel is believed to have the most accurate, reliable, and efficient performance in real-world applications [19].

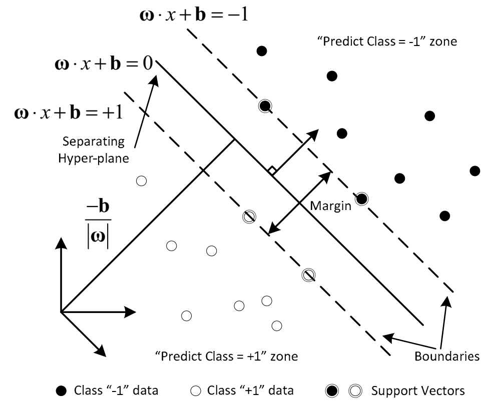

Fig. 4. SVM classification.

B. Multiclass SVM Classification

The fault classification of a drivetrain gearbox is a multiclass problem. Therefore, a multiclass SVM classifier is designed. Choosing an appropriate classification strategy is a critical issue in multiclass classification, and much work has been done on this subject [20]. This work utilizes a One-Against-One (OAO) strategy, taking its advantages in a small number of training samples for each classifier, symmetric data structure, and low computational loads [20].

A multiclass SVM classifier with a RBF kernel is designed to evaluate the condition of the test gearbox. The SVM has 4 inputs, which are the 4 features generated by the adaptive feature extraction scheme described in Section III. The output of the SVM is the code of the fault types as in Table I.

TABLE I
OUTPUT CODING IN PROPOSED MULTICLASS SVM CLASSIFIER

Condition	Health	One-tooth-missing	Two-teeth-missing	Gear crack
Code	0	1	2	3

V. PSO Algorithm to Optimize Multiclass SVM

A. Principles of PSO Algorithm

The PSO algorithm is a population-based stochastic optimization method inspired by the social behavior of bird blocking or fish schooling. In the description of PSO, the swarm is constituted of a certain number of particles moving in the problem hyperspace to search for the global optima iteratively. Each particle has a position vector and a velocity vector for directing its movement. The PSO algorithm is implemented in the following iterative procedure[21], [22].

(i) Initialize a population of particles with random positions and velocities of M dimensions in the problem space;
(ii) Evaluate the fitness function for each particle;
(iii) Compare each particle's fitness value with its previous best fitness $pbest_i$. If the current value is better than $pbest_i$, then set this value as $pbest_i$ and the particle's current position x_i as p_i;
(iv) Identify the particle in the neighborhood with the best fitness value and set this value as $gbest$ and the particle 's position as p_g;
(v) Update the velocity and position of each particle according to the following two equations, respectively;

$$v_i(k) = wv_i(k-1) + c_1 r_1[p_i - x_i(k-1)] + c_2 r_2[p_g - x_i(k-1)] \quad (11)$$

$$x_i(k) = x_i(k-1) + v_i(k) \quad (12)$$

(vi) Repeat steps (ii)-(v) until a stopping criterion, e.g., a sufficiently good fitness or a maximum number of iteration, is met. The final value of p_g is regarded as the optimal solution of the problem.

In (11) and (12), i is the index of the particle and $i = 1, 2, \dots, N$; k is the current number of iteration; w is the inertia weight, which is a positive scalar used to provide better control between exploration and exploitation; r_1 and r_2 are two M-dimensional vectors of random numbers uniformly distributed in $[0, 1]$; c_1 and c_2 are acceleration coefficients determining the magnitude of the random forces that attract each particle toward the individual best p_i and neighborhood best p_g, respectively. In addition, in some applications, the velocity v_i in step (v) is limited to the range $[-V_{max}, V_{max}]$ and the position x_i is limited to the range $[-X_{max}, X_{max}]$.

B. SVM Parameter Optimization using PSO

The PSO method is employed to optimize the parameter setting of the multiclass SVM classifier. The SVM parameters that need to be optimized include the penalty number C and the RBF kernel width γ. Therefore, the dimension of the problem hyperspace is $M = 2$. The PSO algorithm is designed to search for the optimal values of the parameter set $x = [C, \gamma]$ to optimize the SVM classification accuracy, where the fitness function is the SVM classification accuracy. The parameters of the PSO are listed in Table II.

TABLE II
PSO PARAMETER SETTING

Swarm Size (N)	20	c_1	1.5	w	1
Iteration	80	c_2	1.6	$V_{c,max}$	50
$V_{\gamma,max}$	500	C_{max}	100	γ_{max}	1000
C_{min}	0.1	γ_{min}	0.1		

In the multiclass SVM classifier, the values of C and γ should always be positive. Hence, it is important to choose appropriate boundary conditions for position x and velocity v in the PSO algorithm. Normally, C lies in $[0.1, 100]$, and γ lies in $[0.1, 1000]$. In order for the particles to search for the best position adequately within the boundaries, the constraints on the particle velocities are set according to the particle position boundaries as follows: $V_{c,max} = k_1 C_{max}$, $V_{\gamma,max} = k_2 \gamma_{max}$, and $k_1 = k_2 = 0.5$.

VI. EXPERIMENTAL STUDIES

A. Experimental Setup

The proposed method is validated for diagnosis of multiple types of faults in a test drivetrain gearbox. Fig. 5 shows the experimental system setup, which consists of a 300-W PMSM driven by a variable-speed induction motor through two back-to-back connected gearboxes. They are two-stage helical gearboxes with a total gear ratio of 10.57. One gearbox (i.e., the speed reducer) reduces the shaft speed of the induction motor. The second gearbox (i.e., the test gearbox) is used to emulate a drivetrain gearbox in real-system applications. The test gear is mounted at the input shaft of the test gearbox and pretreated by artificially generating various faults which were commonly observed in industrial systems, including one-tooth missing, two-teeth missing, and a gear crack (Fig. 6). One phase stator current of the PMSM is recorded with a sampling rate of 10 kHz for 210 seconds, during which the rotating speed of the PMSM is varied randomly in the range of 297 to 891 RPM. A total number of 1326 data records are collected, including 340 healthy cases, 318 one-tooth-missing cases, 293 two-teeth-missing cases, and 375 gear crack cases.

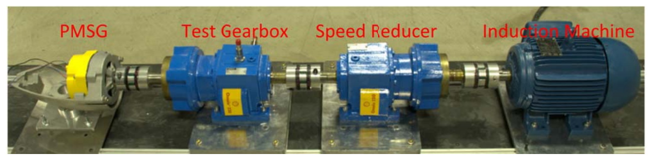

Fig. 5. The experimental system.

(a) (b) (c)

Fig. 6. The test gears with (a) one-tooth missing, (b) two-teeth missing, and (c) a crack.

B. Nonstationary Operating Conditions

As reported in the previous work [5], no physically meaningful signatures can be extracted from the nonstationary stator current signals caused by the variable shaft rotating frequencies using the classical FFT analysis (Fig. 7). The proposed adaptive signal resample algorithm is thus applied to the nonstationary signals acquired from experiments. An example for the PSD spectrum of the resampled PMSM stator current in the healthy case is shown in Fig. 8, where the characteristic frequencies of the gearbox are clearly observed in the PSD spectrum of the resampled current signal. These characteristic frequencies are predicted by the mathematical expression in Section II.

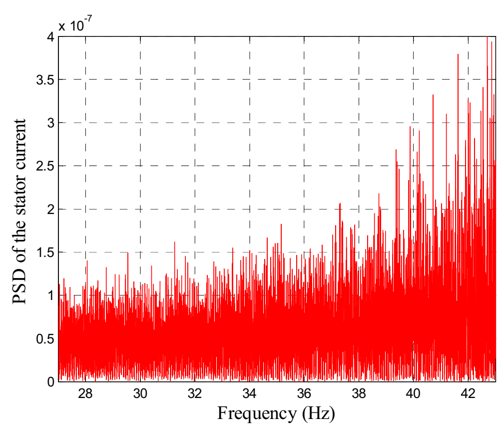

Fig. 7. PSD spectrum of the stator current signal obtained directly from classical FFT analysis.

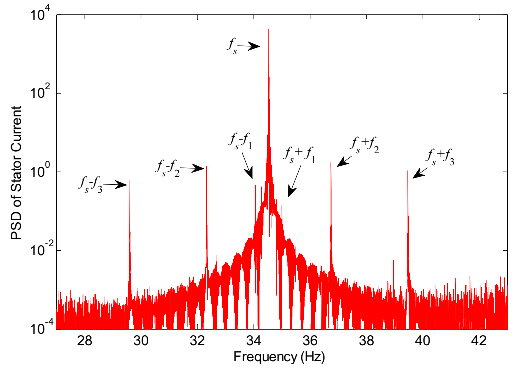

Fig. 8. PSD spectrum of the resampled PMSM stator current in healthy case.

C. Classification Results

The proposed feature extraction scheme is then applied to the PSD spectrum of the resampled PMSM stator current data. Each data record produces one set of four features. The complete dataset have 1326 sets of features.

The PSO algorithm is executed to optimize the parameter setting of the multiclass SVM classifier. The result of the first 20 iterations of the PSO execution is shown in Fig. 9 and the classification results are summarized in Table III. As shown in Fig. 9, the classification accuracy starts with a relatively high value and converges to a steady value after one iteration of the PSO implementation. This indicates that the PSO algorithm can find the optimal parameter setting quickly. This fast convergence is the result of the proper feature extraction in Section III and proper PSO parameter setting in Section V. The results clearly indicate that the multiple types of faults in the drivetrain gearbox are properly classified by the proposed method consistently.

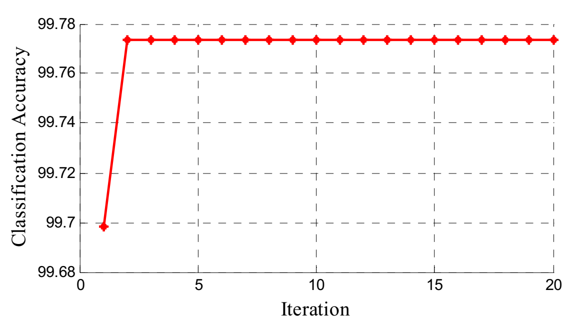

Fig. 9. Accuracy multiclass SVM classification vs. PSO iteration.

TABLE III
PSO OPTIMIZED MULTICLASS SVM CLASSIFICATION RESULTS

	Max	Min	Average	Standard deviation
Accuracy	99.77%	99.69%	99.76%	0.0002

VII. CONCLUSIONS

This paper has presented a novel PSO-optimized multiclass SVM classifier for fault identification of drivetrain gearboxes. An adaptive feature generation algorithm has been proposed to extract the features of the faults from the nonstationary current signals. A multiclass SVM classifier with the RBF kernel has been designed to classify the multiple types of gearbox faults according to the fault features extracted. The PSO algorithm has been adopted to optimize the parameter setting of the multiclass SVM classifier. Experimental studies have been performed for a PMSM-connected drivetrain gearbox with three different faults; and the experimental results have shown that the faults can be effectively classified by using the proposed method with satisfactory classification accuracy.

VIII. REFERENCES

[1] G. Maggetto and J. Van Mierlo, “Electric and electric hybrid vehicle technology: a survey,” in *Proc. IEE Seminar Electric, Hybrid and Fuel Cell Vehicles*, Apr. 2000, pp. 1/1-111.

[2] H. W. Lee, K. Kim, and J. Lee, “Review of maglev train technologies,” *IEEE Trans. Magnetics,* vol. 42, no. 7, pp. 1917-1925, Jun. 2006.

[3] A. Boglietti, A. Cavagnino, A. Tenconi, and S. Vaschetto, “The safety critical electric machines and drives in the more electric aircraft: A survey,” in *Proc. 35th Annu. Conf. of IEEE Ind. Electron. Soc.*, Nov. 2009, pp. 2587-2594.

[4] K. Smolders, H. Long, Y. Fend, and P. Tavner, “Reliability analysis and prediction of wind turbine gearboxes,” in *Proc. European Wind Energy Conf.*, Apr. 2010, pp. 162-165.

[5] D. Lu, X. Gong, and W. Qiao, “Current-based diagnosis for gear tooth breaks in wind turbine gearboxes,” in *Proc. IEEE Energy Convers. Congr. Expo.*, Sept. 2012, pp. 3780-3786.

[6] D. Lu, W. Qiao, X. Gong, and L. Qu, “Current-based fault detection for wind turbine systems via Hilbert-Huang transform,” in *Proc. IEEE Power & Energy Society General Meeting*, Jul. 2013, pp. 1-6.

[7] D. Lu and W. Qiao, “Frequency demodulation-aided condition monitoring for drivetrain gearboxes,” in *Proc. IEEE Transportation Electrification Conference and Exposition*, Jun. 2013, pp. 1-6.

[8] V. N. Vapnik, *The Nature of Statistical Learning Theory*, 2nd ed., New York: Springer, Nov. 2000.

[9] D. Lu and W. Qiao, “Adaptive feature extraction and SVM classification for real-time fault diagnosis of drivetrain gearboxes,” in *Proc. IEEE Energy Convers. Congr. Expo.*, Sept. 2013, pp. 3934-3940.

[10] M. Aly, “Survey on multiclass classification methods,” California Institute of Technology, Tech. Reprot, Nov. 2005.

[11] M. Pal, “Multiclass approaches for support vector machine based land cover classification,” presented at *the 8th Annu. Int. Conf. in the Field of GIS, GPS, Arial Photography, and Remote Sensing*, Feb. 2005.

[12] V. Malathi and N. S. Marimuthu, “Multi-class support vector machine approach for fault classification in power transmission line,” in *Proc. IEEE Int. Conf. Sustainable Energy Technologies*, Nov. 2008, pp. 67-71.

[13] S. R. Gunn, “Support vector machines for classification and regression,” School of Electronics and Computer Science, University of Southampton, Tech. Report, May 1998.

[14] V. Cherkassky and Y. Ma, “Practical selection of SVM parameters and noise estimation for SVM regression,” *Neural Networks,* vol. 17, no. 1, pp. 113-126, Jan. 2004.

[15] X. Yuan and Y. Wang, “Parameter selection of support vector machine for function approximation based on chaos optimization,” *J. Syst. Eng. Electron.,* vol. 19, no. 1, pp. 191-197, Jan. 2008.

[16] C. Hsu, C. Chang, and C. Lin, “A practical guide to support vector classification,” Department of Computer Science, National Taiwan University, Tech. Report, Jul. 2003.

[17] R. R. Schoen and T. G. Habetler, “Effects of time-varying loads on rotor fault detection in induction machines,” *IEEE Trans. Ind. Appl.,* vol. 31, no. 4, pp. 900-906, Jul. 1995.

[18] C. Kar and A. R. Mohanty, “Monitoring gear vibrations through motor current signature analysis and wavelet transform,” *Mech. Syst. Signal Process.,* vol. 20, no. 1, pp. 158-187, Jan. 2006.

[19] C. S. Tyagi, “A comparative study of SVM classifiers and artificial neural networks application for rolling element bearing fault diagnosis using wavelet transform preprocessing,” in *Proc. World Academy of Science, Engineering and Technology*, Sept. 2008, pp. 319-327.

[20] C. Hsu and C. Lin, “A comparison of methods for multiclass support vector machines,” *IEEE Trans. Neural Networks,* vol. 13, no. 2, pp. 415-425, Mar 2002.

[21] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Proc. IEEE Int. Conf. Neural Networks*, Nov. 1995, pp. 1942-1948.

[22] W. Qiao, “Swarm Intelligence for Power System Control and Optimization,” in *Evolutionary Computing for Intelligent Power System Optimization and Control*, 1st ed., Z. Xu and Z. Fan, Eds. New York: Nova Science Publishers, Mar. 2011.