Reliable Fault Diagnosis for Low-Speed Bearings Using Individually Trained Support Vector Machines with Kernel Discriminative Feature Analysis

Myeongsu Kang, Jaeyoung Kim, Jong-Myon Kim, *Member*, *IEEE*, **Andy C. C. Tan, Eric Y. Kim, and Byeong-Keun Choi**

***Abstract*—This paper proposes a highly-reliable fault diagnosis approach for low-speed bearings. The proposed approach first extracts wavelet-based fault features that represent diverse symptoms of multiple low-speed bearing defects. The most useful fault features for diagnosis are then selected by utilizing a genetic algorithm (GA)-based kernel discriminative feature analysis cooperating with one-against-all multi-category support vector machines (OAA MCSVMs). Finally, each support vector machine is individually trained with its own feature vector that includes the most discriminative fault features, offering the highest classification performance. In this study, the effectiveness of the proposed GA-based kernel discriminative feature analysis and the classification ability of individually trained OAA MCSVMs are addressed in terms of average classification accuracy. In addition, the proposed GA-based kernel discriminative feature analysis is compared with four other state-of-the-art feature analysis approaches. Experimental results indicate that the proposed approach is superior to other feature analysis methodologies, yielding an average classification accuracy of 98.06% and 94.49% under rotational speeds of 50 revolutions-per-minute (RPM) and 80 RPM, respectively. Furthermore, the individually trained MCSVMs with their own optimal fault features based on the proposed GA-based kernel discriminative feature analysis outperform the standard OAA MCSVMs, showing an average accuracy of 98.66% and 95.01% for bearings under rotational speeds of 50 RPM and 80 RPM, respectively.**

***Index Terms*—Acoustic emission, fault diagnosis of low-speed bearings, kernel discriminative feature analysis, genetic algorithm, individually trained multi-category support vector machines**

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2012R1A1A2043644). **The authors gratefully acknowledge the use of low-speed fault simulator test data developed within the CRC for Infrastructure and Engineering Asset Management, Queensland University of Technology, established and supported under the Australian Government's Cooperative Research Centres Programme.**

M. Kang is with the School of Electrical, Electronics, and Computer Engineering, University of Ulsan, Ulsan 680-749 South Korea (e-mail: ilmareboy@ulsan.ac.kr).

J. Kim is with the School of Electrical, Electronics, and Computer Engineering, University of Ulsan, Ulsan 680-749 South Korea (e-mail: kim7097@mail.ulsan.ac.kr).

J. -M. Kim is the corresponding author. He is with the Department of IT Convergence, University of Ulsan, Ulsan 680-749 South Korea (e-mail: jmkim07@ulsan.ac.kr).

A. C. C. Tan is with School of chemistry, Physics and Mechanical Engineering, Faculty of Science and Engineering, Queensland University of Technology, Brisbane, Australia (e-mail: a.tan@que.edu.au).

E. Y. Kim is with CMOS NorthParkes Mine, PO Box 995, Parkes, NSW 2870, Australia (e-mail: yonghan.kim@gmail.com).

B. -K. Choi is with the Department of Energy Mechanical Engineering at Gyeongsang National University, Tongyeong City, South Korea (e-mail: bgchoi@gnu.ac.kr).

I. Introduction

LOW-SPEED rotating machines have been widely utilized in industries such as paper mills and wind-turbine power plants [1]. Bearings are the most significant elements in these machines, as they support heavy loads with stationary rotational speeds. Unexpected bearing failures have frequently occurred in recent years because modern variable-speed drives utilize rapid rising voltage pulses and high switching frequencies which can produce current pulses through the bearings, and whose repeated discharges can gradually erode bearing raceways. Such bearing defects can lead to severe motor breakdown and tremendous economic losses, and consequently reliable condition monitoring and bearing defect diagnosis are urgently required to address this issue. In the field of bearing defect diagnosis, vibration analysis has been extensively utilized since it provides the most intrinsic information about diverse bearing failures [2]–[9]. Moreover, motor current signature analysis (MCSA) has been widely exploited for diagnosis of the bearing failures due to the following advantages [10]–[15]: 1) it enables low-cost diagnosis compared to vibration analysis because it does not require any special devices to be installed on the motor and 2) it provides a method for non-intrusive monitoring. To guarantee highly-reliable bearing defect diagnosis, some researchers have also utilized multiple signatures obtained from both vibration and current sensors [16], [17]. Though each of these analyses have shown satisfactory performance identifying multiple bearing defects, their focus has been on separating multiple bearing failures under high rotational speed, ranging from a few hundred to a few thousand revolutions-per-minute (RPM), due to the difficulty of capturing intrinsic information about low-speed bearing defects from weak vibration and current

signals. Acoustic emission (AE) has been an attractive approach to low-speed bearing defect diagnosis, given its ability to capture low-energy signals [1], [18]–[25]. Specifically, Yoshioka et al. and Tandon et al. showed that AE could recognize bearing defects before they appear in the vibration acceleration range or appear on the bearing's surface [26], [27]. This study employs an AE technique for incipient and low-speed bearing fault diagnosis.

Signal processing-based fault diagnosis methods have been widely used in recent years to identify multiple bearing defects such as a crack or spall on the outer or inner raceways of a bearing, involving the following three essential steps: fault extraction, feature analysis to select the most discriminative fault features for diagnosis, and fault classification. Feature calculation is a fundamental step in the mapping of original signals onto the statistical parameters reflecting diverse symptoms of bearing defects, and is performed via time domain analysis [5], [18], [19], frequency domain analysis [16], [17], [21], and time-frequency domain analysis [10], [20], [28]–[46].

Among these analyses, time-frequency domain analysis has garnered increasing interest in research related to capturing intrinsic information about non-stationary bearing defects (frequency information of bearing defects changes over time). Thus, wavelet-based fault features have been widely used for fault diagnosis [39]–[48]; however, a high-dimensional feature vector, which consists of these wavelet-based fault features, can be a primary reason for classification accuracy degradation because there is no guarantee that all of the computed fault features are equally useful for rolling element bearing diagnosis. Thus, fault feature analysis, which is considered as either feature selection or dimensionality reduction of the feature vector, is needed to find the most discriminative fault features in the given feature vector while keeping the intrinsic information about the defects. Several approaches have been introduced for fault feature analysis [49], [50]–[58]. Among various methods for finding useful fault features, component analyses such as principal component analysis (PCA) [1], [2], [51]–[55] and linear discriminant analysis (LDA) [1], [2], [56]–[58] have been widely utilized in fault diagnosis. Although PCA, which is one of the unsupervised analysis methods, is effective for fault feature analysis and the resultant principal components by PCA can provide alternatives to discriminative fault features for diagnosis, this method has a problem preserving the discriminative properties of the data due to the lack of an inter-category separability estimation process.

Unlike PCA-family approaches, LDA, which is one of the supervised analysis techniques that uses the label information of categories, can preserve discriminative information well. Since LDA locates the optimal low-dimensional representation for the high-dimensional feature vector by computing between-category and within-category scatter matrices, LDA-based approaches generally provide better classification results than those obtained via PCA-family approaches [2], [56]. Jin et al. introduced an orthogonal variant of LDA in 2014, called the trace ratio linear discriminant analysis (TR-LDA1) [2], [56], in order to eliminate redundant information from the scatter matrices in LDA [2], [56]. Though the authors in [2], [56] achieved satisfactory performance for bearing failure diagnosis by using TR-LDA1, they extended it to deal with non-Gaussian fault features that can be faced in many fault diagnosis problems. This is because TR-LDA1 was developed with the assumption that fault features follow a Gaussian distribution, which may degrade classification performance. In the extended TR-LDA (TR-LDA2), two new scatter matrices were developed to characterize intra-category compactness and inter-category separability by exploiting intrinsic and penalty graphs. The penalty graph characterizing inter-category separability in TR-LDA2 cannot reflect the neighborhood relationships between different categories, and this can cause classification accuracy degradation. These drawbacks motivated our research to develop a new genetic algorithm (GA)-based kernel discriminative fault feature analysis approach that works well in conjunction with one-against-all multi-category support vector machines (OAA MCSVMs). The proposed method selects the most discriminative fault features in the given feature vector for diagnosis, regardless of fault feature distribution. As classifiers, this study employs OAA MCSVMs, which offer higher classification performance with a limited training data set compared to other classification models such as the artificial neural network [59].

The contributions of this paper are summarized as follows:

- Overhung fans and pumps are widely used in industry, and many contain shafts that are supported by two spherical or cylindrical roller bearings mounted in plummer blocks. In overhung fans and pumps, the drive-end (DE) bearing is slightly loaded, while the non-drive-end (NDE) bearing is highly loaded (the NDE bearing has approximately three times as much load as the DE bearing). This paper provides a useful guideline for determining which bearing, the DE or the NDE bearing, is defective, and which types of bearing defects occur in either the DE bearing or the NDE bearing in overhung fan pumps by carrying out fault diagnosis with multiple bearing defects acquired from various load conditions.
- An efficient GA-based discriminative fault feature analysis approach is proposed for highly-reliable fault diagnosis in low-speed bearings. A GA attempts to locate the optimal combination of fault features for each category (each type of bearing diagnosed in this study) by cooperating with OAA MCSVMs. Likewise, this paper individually trains SVMs by exploiting each optimal feature vector for each SVM structure in order to maximize the classification ability of OAA MCSVMs for diagnosis.
- Multiple bearing defects are acquired under different load conditions and different bearing rotational speeds in this study, and they are used to validate the effectiveness of the proposed GA-based discriminative fault feature analysis as well as the efficacy of individually trained OAA MCSVMs. Experimental results indicate that the proposed fault diagnosis methodology using individually trained OAA MCSVMs with GA-based kernel discriminative feature analysis yields the highest classification accuracy.

The rest of the paper is organized as follows. Section II

describes a test rig for experiments and introduces multiple bearing defects to be diagnosed in this study. Section III presents the proposed diagnosis methodology including GA-based kernel discriminative feature analysis, and Section IV validates the effectiveness of proposed methodologies in terms of classification accuracy. Finally, Section V concludes this paper.

II. A Fault Machinery Simulator and Bearing Defects

Data obtained from a low-speed machinery fault simulator for fault diagnosis of rolling element bearings **developed by CRC-IEAM, Queensland University of Technology (QUT), Australia, was used in this paper,** as shown in Fig. 1(a) **[60]-[62]**. **The test rig can simulate a range of bearing and gear faults with loads applied through a radial mechanism and a hydraulic brake, as shown in Fig. 1(a).** In addition, a cylindrical roller bearing (SKF NF307) was used for the test. At the drive-end of the test rig, the shaft is connected to a reduction gear box (10.1:1) through a coupling. Likewise, a constant radial load can be applied to the drive-end support and is measured by a load cell. To continuously acquire AE signals, a general-purpose AE sensor (PAC R3α) was attached on the top of the bearing housing, as depicted in Fig. 1(b). Detailed specifications of data acquisition system are presented in Table I.

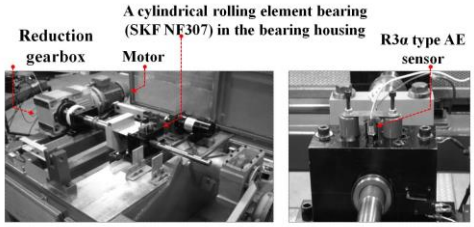

Fig. 1. (a) The test rig for the experiments **[60]** and (b) the location of an AE sensor to record continuous AE signals **[61], [62]**

TABLE I
DETAILED SPECIFICATIONS OF DATA ACQUISITION SYSTEM

PCI-based AE system	•	18-bit 10 MHz A/D conversion
AE sensor	•	Peak sensitivity (V/μbar): -63dB
	•	Operating frequency range: 25 – 530 kHz
	•	Resonant frequency (V/μbar): 140 kHz
	•	Directionality: ±1.5 dB
The operating frequency is set to 500 kHz in this study.		

In order to diagnose various bearing defects, various seeded defects **developed by QUT were used**, as illustrated in Fig. 2. A normal bearing (or a defect-free bearing, NB) is used as a reference case in this study. In total, 12 types of AE signals (six under a 500-N load and another six under a 2-kN load) were acquired from bearings rotating at 50 RPM and 80 RPM, respectively. A diamond cutter bit and an air-speed grinding tool were used to produce cracks and spalls, respectively, on the bearing surface. Table II presents a detailed description of the seeded bearing defects, and 90 1.5-second AE signals sampled at 500 kHz were used for diagnosing each bearing condition in this study.

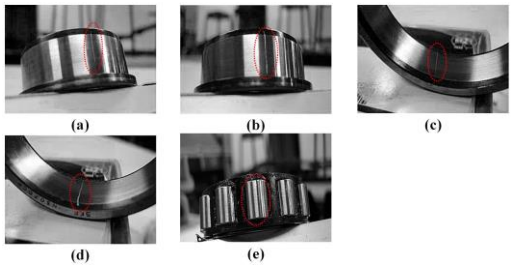

Fig. 2. Various seeded bearing defects **[61], [62]**. (a) Inner-race crack (IRC), (b) inner-race spall (IRS), (c) outer-race crack (ORC), (d) outer-race spall (ORS), and (e) roller medium spall (RMS)

TABLE II
DETAILED DESCRIPTIONS FOR SEEDED BEARING DEFECTS **[61], [62]**

Seeded bearing defects	Width (mm)
Hair-line crack on inner-race (IRC)	0.1
Small-line spall on inner-race (IRS)	0.6
Hair-line crack on outer-race (ORC)	0.1
Small-line spall on outer-race (ORS)	0.7
Medium-line spall on roller (RMS)	1.6

III. Proposed Diagnosis Methodology

The proposed diagnosis methodology for low-speed bearing defects includes feature calculation, GA-based kernel discriminative fault analysis cooperating with OAA MCSVMs, and fault classification. More details about each step are given below.

A. Feature Calculation

To record continuous AE signals for identifying diverse bearing defects, an AE sensor is generally attached to a non-rotating part of the machinery (e.g., bearing housing). Although the bearing housing is the closest element to place an AE sensor, the distance from the source of bearing failures causes severe attenuation in the recorded AE signals. This attenuation can be one of the reasons why intrinsic information about diverse bearing failures exists primarily in mid- and high-frequency bands. To deal with this issue, discrete wavelet transform (DWT) has been extensively employed to analyze failure information inherent in AE signals because of its decomposition ability, which splits a signal into low- and high-frequency bands. In addition, DWT is a promising time-frequency analysis tool while processing non-stationary AE signals **[63], [64].**

Wavelet packet transform (WPT) is more effective for decomposing both mid- and high-frequency information from a signal into both mid- and high-frequency regions rather than DWT. For this reason, WPT is initially performed in this study with a 1.5-second AE signal to extract fault features. According to Yan et al. [48], both relative energy in a wavelet packet node (REWPN) and entropy in a wavelet packet node (EWPN) are effective for revealing the disorder behaviors of a signal, which are generally considered as fault symptoms in fault diagnosis applications. Both are used in this study as fault features for diagnosis. To compute these fault features, a 1.5-second AE signal is decomposed via three-level WPT, and eight wavelet packet nodes are generated, as shown in Fig. 3.

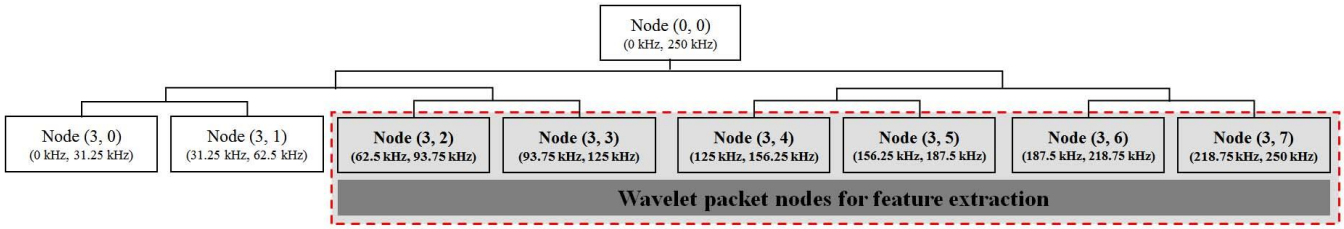

Fig. 3. Three-level WPT on an AE signal

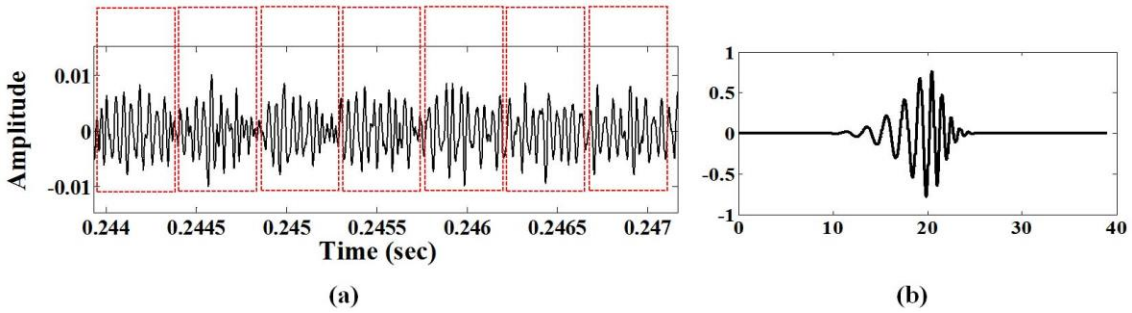

Fig. 4. (a) Waveform of an AE signal and (b) wavelet mother function (db20)

As previously mentioned, intrinsic information revealing bearing defects lies primarily in high-frequency bands, and thus in this study wavelet-based fault features (REWPN and EWPN) are extracted from the last six mid- and high-frequency wavelet packet nodes. Since the choice of a mother wavelet function when performing both DWT and WPT greatly influences their analysis results, it is significant to select an appropriate mother wavelet function in order to obtain a useful description of AE signals. In this study, Daubechies 20 (or db20) is used for signal decomposition because it is one of the best matches to the acquired AE signals, as shown in Fig. 4.

First, REWPN is defined as follows:

$$REWPN(i) = \sum_{j=1}^{K} w_{i,j}^2 \Big/ \sum_{n=1}^{N_{tnode}} \sum_{j=1}^{K} w_{n,j}^2, \quad (1)$$

where N_{tnode} is the total number of wavelet packet nodes considered in this study (N_{tnode} = 6), K is the total number of wavelet coefficients in each wavelet packet node, and $w_{i,j}$ is the jth wavelet coefficient of the ith wavelet packet node.

Second, EWPN is computed as:

$$EWPN(i) = -\sum_{j=1}^{K} p_i(j) \log_2 p_i(j), \quad (2)$$

where $p_i(j) = w_{i,j}^2 \Big/ \sum_{j=1}^{K} w_{i,j}^2$. In total, 12 fault features, including six REWPNs and six EWPNs are used for the diagnosis of low-speed bearing defects in this study.

B. GA-Based Kernel Discriminative Fault Feature Analysis

A genetic algorithm (GA) is a heuristic optimization method based on Darwinian natural selection and genetics in a biological system. The GA is employed in this study for kernel discriminative feature analysis, which finds the most significant fault features in order to identify multiple faults in the rolling element bearings. An optimal feature vector can be represented by a chromosome that is composed of multiple genes, and a gene corresponds to a fault feature (a statistical parameter for describing rolling element bearing defects). Consequently, an $N_{feature}$-bit binary-encoded chromosome (true or false) is used for an optimal feature vector's design, where $N_{feature}$ is the number of features for the purpose of identification of bearing failures in this study ($N_{feature}$=12). For example, if the value of a gene in the $N_{feature}$-bit chromosome is true, this study considers the corresponding feature as a discriminative signature for identifying multiple bearing defects, and vice versa. The GA works with a set of candidate solutions, referred to as a population, and obtains an optimal solution after a series of iterative computations in which the population is a set of chromosomes. The GA generates successive populations of alternative solutions, which can be represented by a chromosome. In this study, an evaluation starts from the population with randomly generated chromosomes. An objective function is utilized to evaluate the quality of a solution (a chromosome). The GA searches for better solutions using a genetic operation, including selection and crossover operations. The selection operation selects superior chromosomes out of the current population to be parents that can generate offspring, and the objective function is used to determine whether or not the chromosomes are superior. The crossover operation is used to select genes from the parent chromosomes and to create new offspring. Since the objective function is utilized in the GA to evaluate the quality of a solution, the design of the objective function is significant. Thus, a proper objective function is proposed in this study to identify an optimal feature vector by cooperating with OAA MCSVMs.

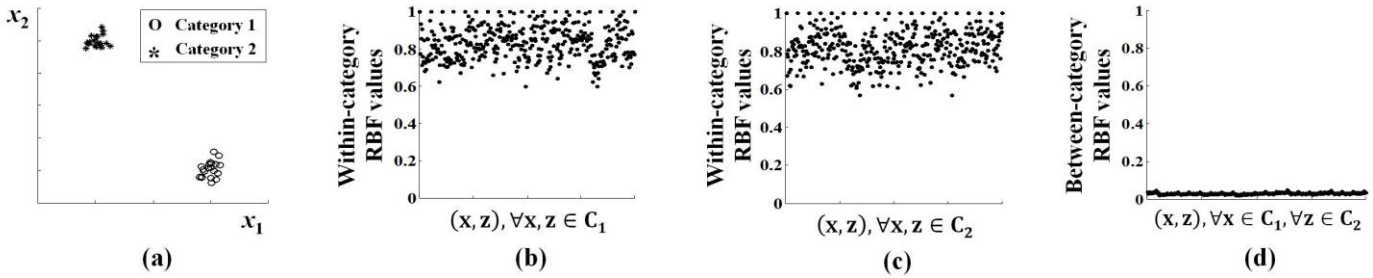

Fig. 5. An example exploring the impact of both within-category RBF values (intra-category compactness) and between-category RBF values (inter-category separability) with two separable categories

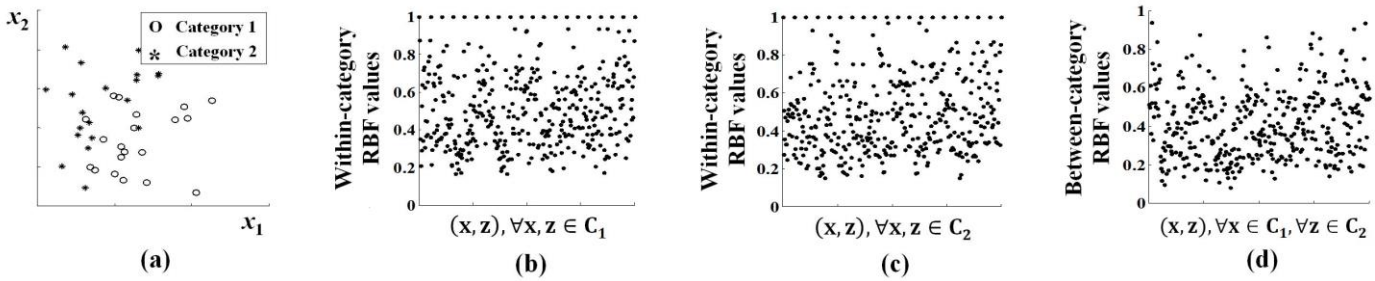

Fig. 6. An example exploring the impact of both within-category RBF values (intra-category compactness) and between-category RBF values (inter-category separability) with two indistinguishable categories

.

An SVM locates a hyperplane between the two categories with the largest margin in the feature space and this hyperplane is used for classifying test samples into one of the two categories. Let $x_i \in R^d,\ \forall i = 1,2,...,n$ and $y_i \in \{-1,+1\},\ \forall i = 1,2,...,n$ be a set of training samples and the corresponding labels, respectively, where n is the total number of samples. To find the optimal hyperplane for separating two categories, the following minimization optimal problem can be solved **[65]**:

$$\arg\min_{w,\xi_i}\left\{\frac{1}{2}w^T w + C\sum_{i=1}^{n}\xi_i\right\},$$
$$\text{subject to } y_i\left(w^T\phi(x_i)+b\right)\geq 1-\xi_i,\ \ \xi_i\geq 0,\ \forall i=1,2,...,n, \tag{3}$$

where w is a normal vector to the hyperplane, b is a constant such that $b/\|w\|$ represents the Euclidean distance between the hyperplane and the origin of the feature space, ϕ is a nonlinear function to map the original feature space into the high-dimensional nonlinear feature space, the ζ_i 's are the slack variables to control the training errors, and C is a penalty variable to tune the generalization capability. According to Aydin et al. **[66]**, the minimization optimal problem above can be written in dual form by applying Lagrange optimization as follows:

$$\arg\max_{\alpha_i}\left\{\sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j \phi(x_i)^T\phi(x_j)\right\},$$
$$\text{subject to } \sum_{i=1}^{n}\alpha_i y_i = 0,\ 0\leq\alpha_i\leq C,\ \forall i=1,2,...,n, \tag{4}$$

where the α_i 's are Lagrange multipliers, and x_i and x_j are any two different samples in the training data set. In addition, $\phi(x_i)\cdot\phi(x_j)$ can be replaced with the kernel function by Mercer's theorem **[67]**, where $\cdot$ is the inner product between two vectors:

$$k(x_i,x_j) = \phi(x_i)\cdot\phi(x_j). \tag{5}$$

As a kernel function for the SVM, the Gaussian radial basis function (RBF) kernel has been extensively used, showing satisfactory performance. The RBF kernel is calculated as follows:

$$k(x_i,x_j) = \exp\left(-\gamma\left\|x_i - x_j\right\|^2\right), \tag{6}$$

where $\gamma = \frac{1}{2\sigma^2}$ and σ is an adjustable parameter to be carefully tuned. If σ is small, the exponential is linear, and the higher-dimensional projection loses its non-linear power. On the other hand, if σ is large, the decision boundary is very sensitive to noise during training due to the lack of regularization.

The RBF kernel measures the similarity between two input samples, and the results of (7) and (8) are defined as a within-category RBF value and a between-category RBF value, respectively:

$$k(x,z) \approx 1,\ \forall x,z \in C_i,\ \forall i = 1,2,...,L, \quad (7)$$

$$k(x,z) \approx 0,\ \forall x \in C_i,\ \forall z \in C_j,\ \forall i,j = 1,2,...,L,\ i \neq j, \quad (8)$$

where C_i is a set of samples in the category i, i = 1, 2, …, L, where L is the number of categories. Figs. 5 and 6 illustrate examples showing the impact of both within-category RBF and between-category RBF values for two sample sets. Within-category RBF values for each category are significantly high in the case of two clearly separable categories, as demonstrated in Fig. 5(a), while between-category RBF values are close to 0. On the other hand, when the two categories are not clearly separable, both within-category RBF and between-category RBF values are widely distributed in the range from 0.2 to 1, as shown in Fig. 6(a). This study utilizes the following two design criteria for the objective function in order to measure the above properties. The first criterion, $RBF_{within}(F_{mat})$, is the mean of within-category RBF values resulting from F_{MAT}, which is an $L \times M \times N$ matrix, where L is the number of categories (or the number of bearing failure types and a defect-free bearing), M is the total number of samples for a category in a training dataset (i.e., this study utilizes the same number of samples for each category), and N is the number of fault signatures that are sorted out from the initially produced feature vector by the GA:

$$RBF_{within}(F_{mat}) = \frac{1}{L \times M} \sum_{i=1}^{L} \sum_{j=1}^{M} \sum_{k=1}^{M} k\left(F_{mat}(i,j,:), F_{mat}(i,k,:)\right). \quad (9)$$

The second criterion, $RBF_{between}(F_{mat})$, is the mean of between-category RBF values resulting from F_{mat}, which is calculated by:

$$RBF_{between}(F_{mat}) = \frac{1}{L \times (L-1) \times M^2} \sum_{i=1}^{L} \sum_{\substack{j=1 \\ j \neq i}}^{L} \sum_{k=1}^{M} \sum_{l=1}^{M} k\left(F_{mat}(i,k,:), F_{mat}(j,l,:)\right). \quad (10)$$

According to (7) and (8), a set of fault signatures providing both large and small values of $RBF_{within}(F_{mat})$ and $RBF_{between}(F_{mat})$, respectively, is useful for maximizing the inter-category separability and minimizing the intra-category compactness. Hence, this study evaluates the quality of solutions by designing the objective function with these two criteria to yield the highest classification performance, which is defined as follows:

$$Obj(F_{mat}) = 1 - RBF_{within}(F_{mat}) + RBF_{between}(F). \quad (11)$$

A lower objective value corresponds to higher intra-category compactness and inter-category separability. Finally, the GA finds the most discriminative fault features maximizing both intra-category and inter-category separability in the given feature vector. In other words, this paper explores an optimal feature vector that yields the lowest objective value.

C. Fault Classification

An SVM discriminates test samples into one of two categories, and consequently multi-category SVMs (MCSVMs) are utilized to identify multiple bearing defects in this study. The following three approaches can be considered in the design of MCSVMs: one-against-all (OAA), one-against-one (OAO), and one-acyclic-graph (OAG) [65]. The OAA method, one of the most popular and simplest techniques for multi-category classifiers, is employed in this study. In the OAA approach, each SVM structure separates one category from the others, and the final decision can be made by selecting an SVM structure that yields the highest output value. In order to design OAA MCSVMs, each SVM structure is separately evaluated to achieve the maximum classification accuracy for its own category. All of the SVM structures then cooperate to make a final decision.

IV. Experimental Results

A. Training and Test Data Configuration

To estimate the generalized classification accuracy, k-fold cross validation (k-cv) is employed in this study [68]. In k-cv, a set of feature vectors needed for evaluating classification performance of the proposed methodology is randomly split into k mutual folds, denoted as A_1, A_2, …, A_k. Then, classification accuracy is estimated k times by training and testing OAA MCSVMs. In other words, fold A_j is treated as the training set while the remaining folds are exploited for testing OAA MCSVMs at jth iteration in k-cv. For the purpose of the cross validation scheme, this study sets k to 3. More specifically, 90 initially created feature vectors are divided into three mutual folds (each fold includes 30 randomly-divided feature vectors for each bearing condition), and each of three mutual folds (30 feature vectors for each bearing condition) is reserved as the training set. The remaining folds (60 feature vectors for each bearing condition) are used as the testing set. Consequently, classification accuracy is computed by testing OAA MCSVMs in the remaining folds; the final classification accuracy is the average value of the accuracies attained in each fold. Likewise, the GA-based kernel discriminative feature analysis is performed during training OAA MCSVMs.

B. Performance Evaluation

In this paper, a new GA-based kernel discriminative feature analysis is proposed, which cooperates with OAA MCSVMs to find the most significant fault features for fault diagnosis in low-speed bearings. To validate the effectiveness of the proposed feature analysis scheme, this study compares classification performance between the proposed method and four other state-of-the-art feature analysis approaches, such as PCA, independent component analysis (ICA), TR-LDA1, and TR-LDA2. An optimal feature vector configuration based on these component analyses is performed by computing the eigenvalues of a covariance matrix and then selecting n components with the first n highest eigenvalues. Despite the fact that component analysis-based approaches show satisfactory performance in the area of bearing defect diagnosis, no general consensus has been reached on the number of components offering the highest classification accuracy.

Consequently, it is necessary to explore the impacts of principal, independent, or discriminative components in terms of classification performance, because the n principal, independent, or discriminative components providing the

highest classification accuracy are generally considered optimal fault features for diagnosis. In this study, the training data in *k*-cv is partitioned into *k* sub-folds, and classification performance is measured using one sub-fold for training, and the remaining sub-folds for testing OAA MCSVMs. To obtain precise classification results, *k*-cv is performed 10 times in this study. Fig. 7 illustrates the average classification accuracy with varied numbers of components under different bearing rotational speeds. Classification accuracy in this study is computed as follows:

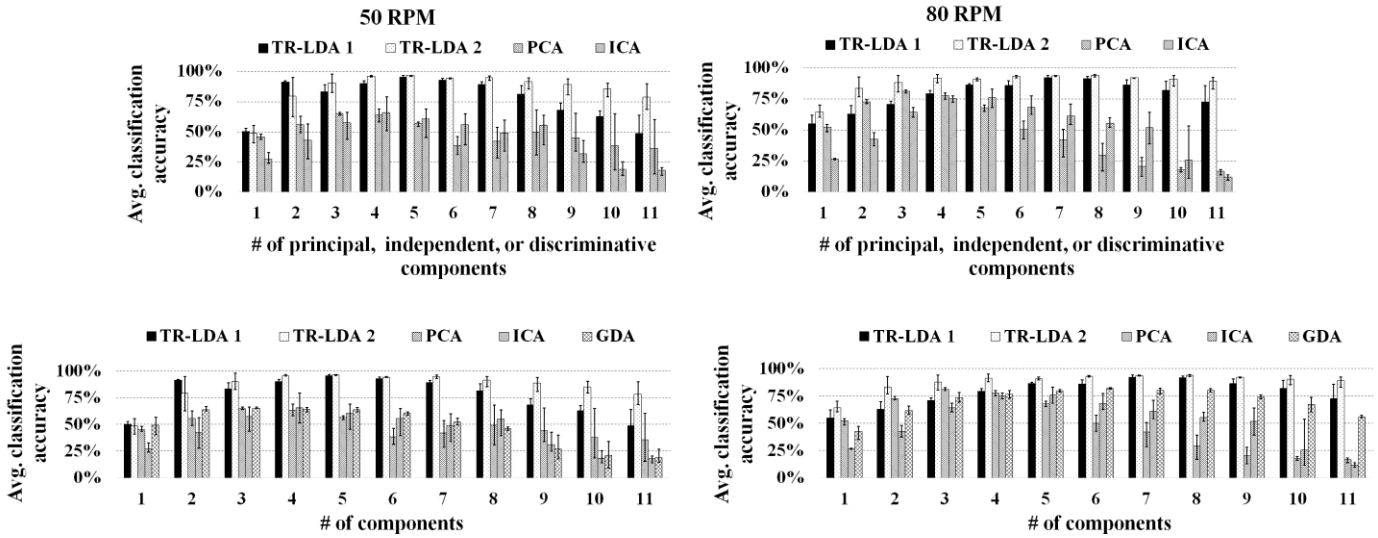

Fig. 7. Average classification accuracies with various numbers of principal, independent, or discriminant components under rotational bearing speeds of 50 RPM and 80 RPM

TABLE III
AVERAGE TPRS AND CLASSIFICATION ACCURACIES FOR MULTIPLE LOW-SPEED BEARING DEFECTS AT A ROTATIONAL SPEED OF 50 RPM (UNIT: %)

	Average TPR per category under 500-N load condition (standard deviation)						Average TPR per category under 2-kN load condition (standard deviation)						Avg. $C_{accuracy}$ (standard deviation)
	IRC1	IRS1	NB1	ORC1	ORS1	RMS1	IRC2	IRS2	NB2	ORC2	ORS2	RMS2	
PCA	29.84 (5.20)	46.22 (4.83)	86.77 (6.58)	35.22 (3.67)	93.84 (2.11)	77.61 (3.35)	39.60 (4.01)	83.51 (3.84)	90.16 (1.62)	64.60 (5.21)	97.94 (2.29)	67.11 (5.14)	67.70 (3.99)
ICA	80.12 (6.17)	83.45 (6.08)	48.29 (11.8)	89.26 (4.25)	76.82 (9.12)	67.00 (10.51)	83.17 (6.91)	82.22 (9.28)	30.26 (19.53)	76.52 (9.27)	73.15 (10.51)	55.61 (9.44)	70.49 (9.41)
TR-LDA1	96.12 (1.83)	96.11 (3.81)	88.05 (3.73)	95.67 (5.57)	98.73 (0.38)	96.11 (3.57)	96.11 (1.87)	99.70 (0.32)	96.34 (1.47)	92.22 (4.97)	100.00 (0.00)	90.55 (2.46)	95.48 (2.50)
TR-LDA2	95.05 (5.85)	94.89 (3.54)	90.33 (4.17)	95.40 (3.64)	98.34 (1.43)	93.49 (4.57)	95.66 (6.32)	97.67 (3.59)	96.23 (1.72)	95.68 (2.86)	99.72 (0.89)	88.04 (4.74)	95.04 (3.61)
Proposed	98.17 (0.70)	98.72 (0.29)	93.84 (1.87)	98.90 (0.00)	99.26 (0.36)	98.94 (0.31)	99.60 (0.60)	99.76 (0.31)	96.17 (1.00)	97.23 (1.54)	100.00 (0.00)	96.17 (1.54)	98.06 (0.71)

$$C_{accuracy} = \frac{\sum_{L} N_{TP}}{N_{samples}} \times 100\ (\%), \tag{12}$$

where L is the number of categories (L=12 in this study), N_{TP} is the number of true positives (*TP*), defined as the total number of faults in category *i* that are correctly classified as category *i*, and $N_{samples}$ is the total number of samples used to evaluate the performance of the proposed bearing failure diagnosis scheme. In this study, the optimal feature vectors are used as inputs for standard OAA MCSVMs and classification performance is measured for 10 *k*-cv instances. Likewise, the selection of optimal pairs of (C, σ) in OAA MCSVMs is significant, as mentioned in Section III.B. A grid search algorithm is employed to determine the best combinations of these parameters per SVM structure in terms of classification performance. The grid search algorithm trains OAA MCSVMs with a pair (C, σ) in the cross-product of the following two sets and evaluates their performances: $C \in \left\{2^{-5}, 2^{-3}, 2^{-1}, 2, 2^{3}, 2^{5}, 2^{7}, 2^{9}, 2^{11}, 2^{13}, 2^{15}\right\}$ and $\sigma \in \left\{2^{-2}, 2^{-1}, 1, 2, 2^{2}, 2^{3}, 2^{4}, 2^{5}, 2^{6}, 2^{7}\right\}$. Finally, the grid search algorithm outputs the best combinations of these parameters, yielding the highest classification performance. In this study, classification accuracy indicates the ability to diagnose an entire category which is defined as 10 bearing defects and two defect-free bearings. Thus, it is necessary to use another performance evaluation index to indicate classification performance for each category. Specifically, this paper utilizes a true positive rate (TPR), which is defined as follows:

$$TPR = \frac{N_{TP}}{N_{TP} + N_{FN}} \times 100\ (\%), \tag{13}$$

where N_{FN} is the number of false negatives (*FN*), defined as the

number of failures in category *i* that are not classified as category *i*. According to Tables III and IV, the proposed fault diagnosis approach using the GA-based kernel discriminative feature analysis yields higher average classification accuracies (Avg. $C_{accuracy}$ in Tables III and IV) than those with four other state-of-the-art feature analysis methods.

TABLE IV
AVERAGE TPRS AND CLASSIFICATION ACCURACIES FOR MULTIPLE LOW-SPEED BEARING DEFECTS AT A ROTATIONAL SPEED OF 80 RPM (UNIT: %)

	Average TPR per category under 500-N load condition (standard deviation)						Average TPR per category under 2-kN load condition (standard deviation)						Avg. $C_{accuracy}$ (standard deviation)
	IRC1	IRS1	NB1	ORC1	ORS1	RMS1	IRC2	IRS2	NB2	ORC2	ORS2	RMS2	
PCA	76.33 (3.96)	57.16 (4.63)	92.93 (1.09)	88.50 (1.86)	94.01 (2.08)	93.34 (1.36)	62.11 (4.88)	47.94 (4.74)	93.45 (2.53)	93.72 (1.99)	99.15 (0.46)	91.06 (4.81)	82.48 (2.87)
ICA	77.95 (6.26)	69.01 (7.14)	88.00 (5.35)	92.22 (3.89)	92.61 (2.61)	79.77 (10.58)	87.54 (3.09)	64.45 (4.47)	87.61 (3.93)	82.06 (4.86)	93.88 (3.43)	71.83 (5.68)	82.24 (5.11)
TR-LDA1	97.39 (0.52)	69.72 (4.91)	98.50 (0.58)	98.38 (1.05)	99.54 (0.68)	95.89 (2.29)	79.24 (5.26)	70.61 (6.30)	99.38 (0.72)	99.31 (0.73)	99.94 (0.19)	90.16 (1.71)	91.51 (2.08)
TR-LDA2	94.71 (2.17)	80.95 (4.50)	94.94 (2.89)	98.20 (1.06)	98.05 (1.18)	92.44 (3.21)	92.81 (3.36)	77.22 (4.49)	96.45 (2.35)	97.22 (1.55)	99.77 (0.55)	83.94 (4.84)	92.24 (2.68)
Proposed	96.98 (1.12)	79.83 (4.50)	97.88 (0.67)	97.46 (1.13)	99.77 (0.55)	97.88 (1.18)	95.78 (1.43)	78.33 (3.34)	97.83 (1.06)	99.77 (0.39)	99.77 (0.55)	92.57 (2.65)	94.49 (1.55)

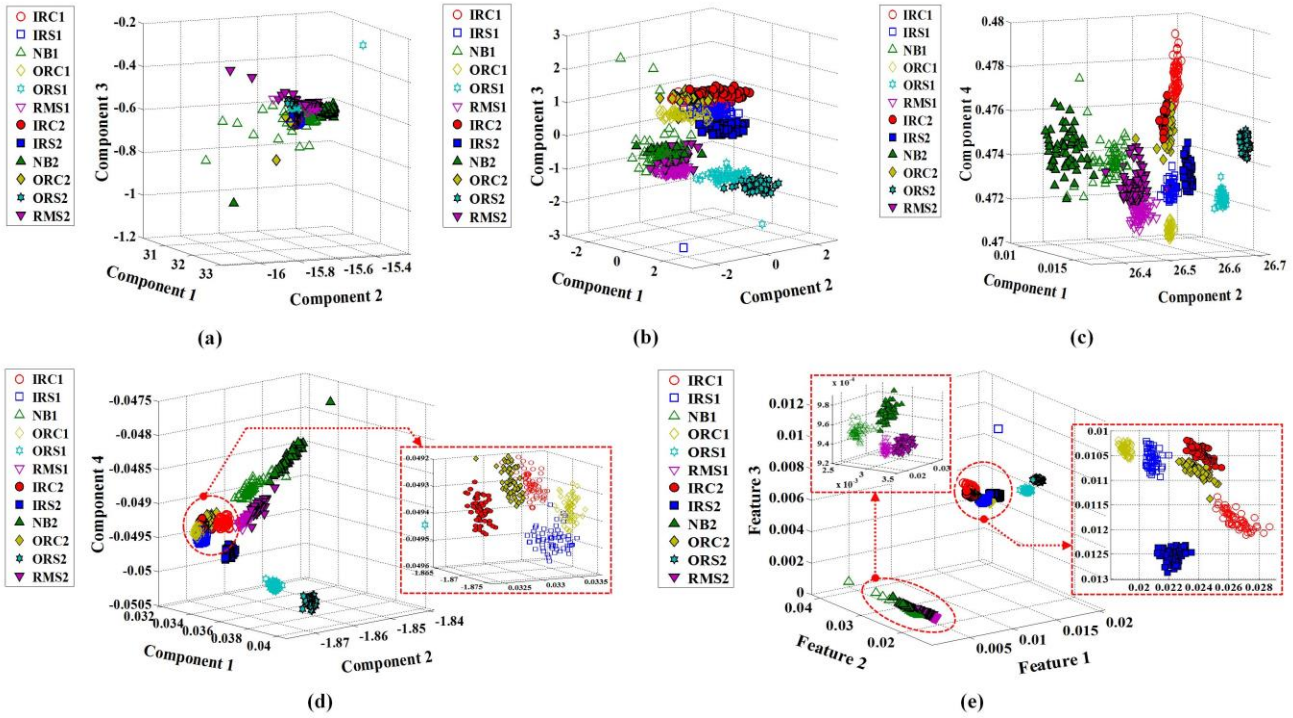

Fig. 8. 3D visualization results of the most discriminative fault features via (a) PCA, (b) ICA, (c) TR-LDA1, (d) TR-LDA2, and (e) the proposed GA-based kernel discriminative feature analysis, where features 1 to 3 are REWPNs computed in the 3rd, 4th, and 5th wavelet packet node after performing the three level WPT using the Daubechies 20-tap decomposition filter. These features are computed using bearing defects at 50 RPM.

In addition to the quantitative evaluation, we present a qualitative evaluation of the proposed approach while comparing it with other approaches in the three-dimensional visualization results of the most discriminative features, as shown in Fig. 8. Unsupervised approaches, such as PCA and ICA, achieve lower average TPRs and classification accuracies than those based on supervised methods (TR-LDAs and the proposed GA-based kernel discriminative feature analysis). The inability to measure inter-category separability is the main reason why these two unsupervised approaches cannot preserve discriminant properties of bearing defects, resulting in unclear and overlapped boundaries among different categories, as demonstrated in Fig. 8. In other words, supervised methodologies yield significantly higher average TPRs and classification accuracies than unsupervised approaches. Specifically, TR-LDAs exhibit satisfactory performance by exploiting intra-category compactness and inter-category separability information, as mentioned in Section I. However, TR-LDAs misidentify bearing defects between NB1 (NB under a 500-N load) and NB2 (NB under a 2-kN load), between

RMS1 (RMS under a 500-N load) and RMS2 (RMS under a 2-kN load). As shown in Fig. 8, the boundaries among these categories heavily overlap and are unclear, resulting in degraded classification performance. Overall, the proposed GA-based kernel discriminative feature analysis approach outperforms these methods because bearing defects of the same category are closely agglomerated, while those belonging to different categories are obviously separated.

TABLE V
CLASSIFICATION PERFORMANCE COMPARISON BETWEEN STANDARD OAA MCSVMS AND INDIVIDUALLY-TRAINED OAA MCSVMS USING THE PROPOSED GA-BASED KERNEL DISCRIMINATIVE FEATURE ANALYSIS FOR LOW-SPEED BEARING FAULT DIAGNOSIS (UNIT: %)

	RPM	Average TPR per category under 500-N load condition (standard deviation)						Average TPR per category under 2-kN load condition (standard deviation)						Avg. $C_{accuracy}$ (standard deviation)
		IRC1	IRS1	NB1	ORC1	ORS1	RMS1	IRC2	IRS2	NB2	ORC2	ORS2	RMS2	
Standard OAA MCSVMs	50	98.17 (0.70)	98.72 (0.29)	93.84 (1.87)	98.90 (0.00)	99.26 (0.36)	98.94 (0.31)	99.60 (0.60)	99.76 (0.31)	96.17 (1.00)	97.23 (1.54)	100.00 (0.00)	96.17 (1.54)	98.06 (0.71)
	80	96.98 (1.12)	79.83 (4.50)	97.88 (0.67)	97.46 (1.13)	99.77 (0.55)	97.88 (1.18)	95.78 (1.43)	78.33 (3.34)	97.83 (1.06)	99.77 (0.39)	99.77 (0.55)	92.57 (2.65)	94.49 (1.55)
Individually trained OAA MCSVMs	50	98.16 (0.74)	97.22 (1.42)	96.78 (0.67)	98.90 (0.00)	100.00 (0.00)	99.32 (0.43)	99.33 (0.89)	99.70 (0.32)	97.93 (0.45)	97.00 (1.77)	100.00 (0.00)	99.54 (0.58)	98.66 (0.61)
	80	97.54 (1.92)	87.84 (4.12)	97.60 (0.85)	95.61 (4.35)	99.37 (0.68)	99.94 (0.19)	86.11 (6.30)	82.61 (2.97)	98.18 (0.65)	99.76 (0.31)	100.00 (0.00)	95.50 (2.11)	95.01 (2.04)

TABLE VI
CONFUSION MATRIX FOR SHOWING CLASSIFICATION RESULTS UNDER ROTATIONAL SPEED OF 80 RPM

	IRC1	IRS1	NB1	ORC1	ORS1	RMS1	IRC2	IRS2	NB2	ORC2	ORS2	RMS2
IRC1	1756	0	0	0	0	0	0	0	0	0	0	1
IRS1	27	1581	0	17	5	0	239	224	0	0	0	0
NB1	0	8	1757	0	0	0	0	5	8	0	0	15
ORC1	1	2	0	1721	0	0	0	0	0	1	0	0
ORS1	0	0	0	0	1789	0	0	1	21	0	0	1
RMS1	0	3	4	0	0	1799	0	7	0	0	0	43
IRC2	0	57	0	0	0	0	1550	41	0	0	0	0
IRS2	12	139	0	57	5	0	11	1487	1	2	0	2
NB2	0	3	23	0	0	0	0	7	1767	0	0	3
ORC2	3	5	3	5	0	0	0	17	0	1796	0	15
ORS2	1	0	0	0	1	0	0	3	1	1	1800	1
RMS2	0	2	13	0	0	1	0	8	2	0	0	1719

C. Standard OAA MCSVMs vs. Individually-Trained OAA MCSVMs

Though OAA MCSVMs with the proposed GA-based kernel discriminative feature analysis generally yield high classification accuracy, classification accuracy degradation can occur, because they share the same fault features for diagnosis. To improve the classification performance of standard OAA MCSVMs utilizing the same fault features, this study trains OAA MCSVMs with an individual feature vector involving the most effective fault features for each SVM structure using the proposed GA-based kernel discriminative feature analysis. To do this, the objective function in (11) should be modified to configure an individual feature vector for SVM_i, where SVM_i works for category i:

$$Obj\left(i,F_{mat}\right)=\max_{j}\left(1-\frac{\left(RBF_{within}\left(i,F_{mat}\right)+RBF_{within}\left(j,F_{mat}\right)\right)}{2}+RBF_{between}\left(i,j,F_{mat}\right)\right),\quad \forall j=1,2,...,L,\ i\neq j, \tag{14}$$

where $RBF_{within}\left(i,F_{mat}\right)=\frac{1}{M}\sum_{k=1}^{M}\sum_{l=1}^{M}k\left(F_{mat}\left(i,k,:\right),F_{mat}\left(i,l,:\right)\right)$ is the mean of the within-category RBF values for category i, and

$RBF_{between}\left(i,j,F_{mat}\right)=\frac{1}{M^2}\sum_{k=1}^{M}\sum_{l=1}^{M}k\left(F_{mat}\left(i,k,:\right),F_{mat}\left(i,l,:\right)\right)$ is the mean of the between-category RBF values between category i and category j. As a result, the proposed objective function maximizes the inter-category separability between category i and the most indistinguishable category. Table V shows that individually trained OAA MCSVMs outperform standard OAA MCSVMs, yielding average classification accuracies of 98.66% and 95.01% under bearing rotational speeds of 50 RPM and 80 RPM, respectively. Although the proposed individually trained OAA MCSVMs-based approach achieves satisfactory classification performance using an optimal combination of discriminative fault features for each category, the bearing defects are not sufficiently separated between IRS1 (IRS under a 500-N load) and IRS2 (IRS under a 2-kN load) under a bearing rotational speed of 80 RPM, as shown in Table VI. Fig. 9 illustrates 3D visualization results of the three most discriminative fault features selected by the proposed GA-based kernel discriminative fault feature analysis for identifying IRS1 and IRS2. Specifically, an SVM structure uses REWPNs (computed in the 3rd and 7th wavelet packet nodes) and EWPN (computed in the 3rd wavelet packet node) in order to identify IRS1, while another SVM structure utilizes REWPNs (computed in the 3rd and 7th wavelet packet nodes) and EWPN (computed in the 4th wavelet packet node) for diagnosing IRS2. As depicted in Fig. 9, the selected discriminative fault features for identifying IRS1 and IRS2 are agglomerated, resulting in misclassification between IRS1 and IRS2. This is highly correlated to signal attenuation. In general, as the AE signals are captured for inner-race bearing defects, signal attenuation occurs more rapidly due to the greater

distance of this source than for other defects considered in this study. This leads to signal attenuation which has an influence on the capture of intrinsic IRS information. Furthermore, the background noise level increases with the rotational speed, adding to the difficulty in extracting discriminative fault features for identifying inner-race bearing defects.

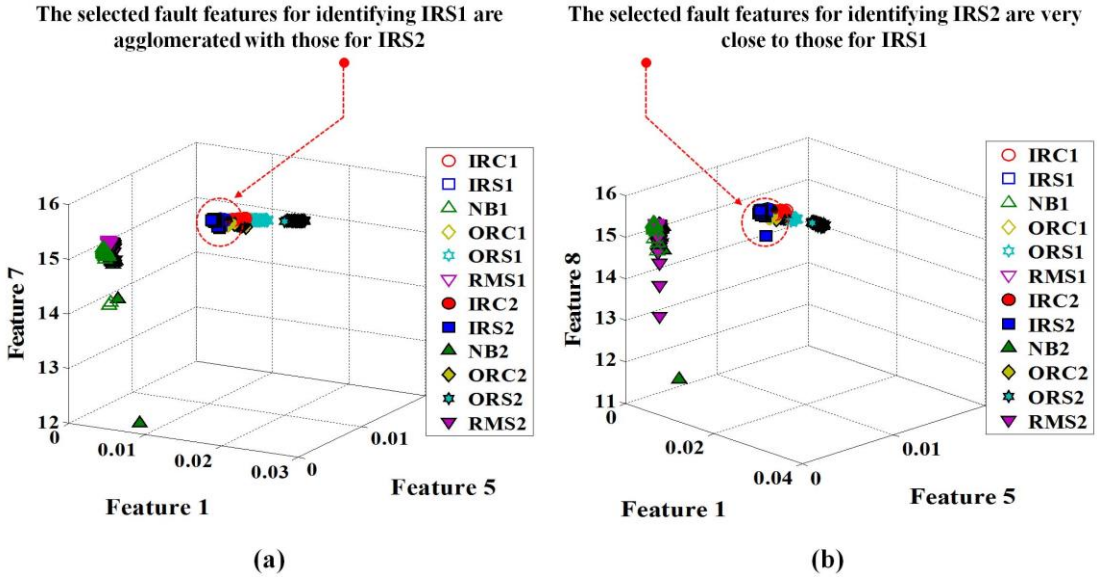

Fig. 9. 3D visualization results of the three most discriminative fault features via the proposed GA-based kernel discriminative feature analysis for identifying (a) IRS1 and (b) IRS2. These features are computed using bearing defects at 80 RPM.

V. Conclusions

This paper proposed a reliable fault diagnosis methodology for low-speed bearings, composed of feature calculation, useful feature selection of bearing defects with the proposed GA-based kernel discriminative feature analysis, and fault classification. This paper first computed wavelet-based fault features, relative wavelet packet energy and wavelet packet node entropy, to represent diverse symptoms of bearing defects. The most useful fault features for each category were then selected in the given feature vector and individually trained OAA MCSVMs were utilized to maximize the classification performance of standard OAA MCSVMs. Finally, multiple low-speed bearing defects were identified by employing the individually trained OAA MCSVMs. Experimental results indicated that the proposed fault diagnosis methodology achieves the highest classification accuracy of 98.66% and 95.01% under bearing rotational speeds of 50 RPM and 80 RPM, respectively.

References

[1] A. Widodo, B. -S. Yang, E. Y. Kim, A. C. C. Tan, and J. Mathew, "Fault Diagnosis of Low Speed Bearing Based on Acoustic Emission Signal and Multi-Class Relevance Vector Machine," *Nondestruct. Test. Eva.*, vol. 24, no. 4, pp. 313–328, 2009.

[2] X. Jin, M. Zhao, T. W. S. Chow, and M. Pecht, "Motor Bearing Fault Diagnosis Using Trace Ratio Linear Discriminant Analysis," *IEEE Trans. Ind. Electron.*, vol. 61, no. 6, pp. 2441–2451, 2014.

[3] J. Seshadrinath and B. Singh, "Investigation of Vibration Signatures for Multiple Fault Diagnosis in Variable Frequency Drives Using Complex Wavelets," *IEEE Trans. Power Electron.*, vol. 29, no. 2, pp. 936–945, 2014.

[4] M. Cococcioni, B. Lazzerini, and S. L. Volpi, "Robust Diagnosis of Rolling Element Bearings Based on Classification Techniques," *IEEE Trans. Ind. Informat.*, vol. 9, no. 4, pp. 2256–2263, 2013.

[5] M. D. Prieto, G. Cirrincione, A. G. Espinosa, J. A. Ortega, and H. Henao, "Bearing Fault Detection by a Novel Condition-Monitoring Scheme Based on Statistical-Time Features and Neural Networks," *IEEE Trans. Ind. Electron.*, vol. 30, no. 8, pp. 3398–3407, 2013.

[6] I. Bediaga, X. Mendizabal, A. Arnaiz, and J. Munoa, "Ball Bearing Damage Detection Using Traditional Signal Processing Algorithms," *IEEE Instrum. Meas. Magz.*, vol. 16, no. 2, pp. 20–25, 2013.

[7] I. Bediaga, X. Mendizabal, I. Etxaniz, and J. Munoa, "An Integrated System for Machine Tool Spindle Head Ball Bearing Fault Detection and Diagnosis," *IEEE Instrum. Meas. Magz.*, vol. 16, no. 2, pp. 42–47, 2013.

[8] A. S. Raj and N. Murali, "Early Classification of Bearing Faults Using Morphological Operators and Fuzzy Inference," *IEEE Trans. Ind. Electron.*, vol. 60, no. 2, pp. 567–574, 2013.

[9] J. Liu, W. Wang, and F. Golnaraghi, "An Extended Wavelet Spectrum for Bearing Fault Diagnostics," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 12, pp. 2801–2812, 2008.

[10] E. C. C. Lau and H. W. Ngan, "Detection of Motor Bearing Outer Raceway Defect by Wavelet Packet Transformed Motor Current Signature Analysis," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 10, pp. 2683–2690, 2010.

[11] L. Frosini and E. Bassi, "Stator Current and Motor Efficiency as Indicators for Different Types of Bearing Faults in Induction Motors," *IEEE Trans. Ind. Electron.*, vol. 57, no. 1, pp. 244–251, 2010.

[12] W. Zhou, B. Lu, T. G. Habetler, and R. G. Harley, "Incipient Bearing Fault Detection via Motor Stator Current Noise Cancellation Using Wiener Filter," *IEEE Trans. Ind. Appl.*, vol. 45, no. 4, pp. 1309–1317, 2009.

[13] A. Ibrahim, M. E. Badaoui, F. Guillet, and F. Bonnardot, "A New Bearing Fault Detection Method in Induction Machines Based on Instantaneous Power Factor," *IEEE Trans. Ind. Electron.*, vol. 55, no. 12, pp. 4252–4259, 2008.

[14] W. Zhou, T. G. Habetler, and R. G. Harley, "Bearing Fault Detection via Stator Current Noise Cancellation and Statistical Control," *IEEE Trans. Ind. Electron.*, vol. 55, no. 12, pp. 4260–4269, 2008.

[15] I. Y. Onel and M. E. H. Benbouzid, "Induction Motor Bearing Failure Detection and Diagnosis: Park and Concordia Transform Approaches Comparative Study," *IEEE/ASME Trans. Mechatronics*, vol. 13, no. 2, pp. 257–262, 2008.

[16] A. Soualhi, G. Clerc, and H. Razik, "Detection and Diagnosis of Faults in Induction Motor Using an Improved Artificial Ant Clustering Technique," *IEEE Trans. Ind. Electron*., vol. 60, no. 9, pp. 4053–4062, 2013.

[17] F. Immovilli, A. Bellini, R. Rubini, and C. Tassoni, "Diagnosis of Bearing Faults in Induction Machines by Vibration or Current Signals: A

Critical Comparison," *IEEE Trans. Ind. Appl.*, vol. 46, no. 4, pp. 1350–1359, 2010.

[18] D. H. Pandya, S. H. Upadhyay, and S. P. Harsha, "Fault Diagnosis of Rolling Element Bearing with Intrinsic Mode Function of Acoustic Emission Data Using APF-KNN," *Expert Syst. Appl.*, vol. 40, no. 10, pp. 4137–4145, 2013.

[19] S. A. Niknam, V. Songmene, and Y. H. J. Au, "The Use of Acoustic Emission Information to Distinguish Between Dry and Lubricated Rolling Element Bearings in Low-Speed Rotating Machines," *Int. J. Adv. Manuf. Technol.*, vol. 69, no. 9–12, pp. 2679–2689, 2013.

[20] W. Caesarendra, P. B. Kosasih, A. K. Tieu, C. A. S. Moodie, and B. -K. Choi, "Condition Monitoring of Naturally Damaged Slow Speed Slewing Bearing Based on Ensemble Empirical Mode Decomposition," *J. Mech. Sci. Technol.*, vol. 27, no. 8, pp. 2253–2262, 2013.

[21] B. Eftekharnejad, M. R. Carrasco, B. Charnley, and D. Mba, "The Application of Spectral Kurtosis on Acoustic Emission and Vibrations from a Defective Bearing," *Mech. Syst. Signal Process.*, vol. 25, no. 1, pp. 266–284, 2011.

[22] M. Elforjani and D. Mba, "Condition Monitoring of Slow-Speed Shafts and Bearings with Acoustic Emission," *Int. J. Exp. Mech.*, vol. 47, no. 2, pp. 350–363, 2011.

[23] M. Elforjani and D. Mba, "Accelerated Natural Fault Diagnosis in Slow Speed Bearings with Acoustic Emission," *Eng. Fract. Mech.*, vol. 77, no. 1, pp. 112–127, 2010.

[24] A. Widodo, E. Y. Kim, J. -D. Son, B. -S. Yang, A. C. C. Tan, D. -S. Gu, B. -K. Choi, and J. Mathew, "Fault Diagnosis of Low Speed Bearing Based on Relevance Vector Machine and Support Vector Machine," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 7252–7261, 2009.

[25] M. Elforjani and D. Mba, "Natural Mechanical Degradation Measurements in Slow Speed Bearings," *Eng. Fail. Anal.*, vol. 16, no. 1, pp. 521–532, 2009.

[26] T. Yoshioka and T. Fujiwara, "Application of Acoustic Emission Technique to Detection of Rolling Bearing Failure," *Am. Soc. Mech. Eng.*, vol. 14, pp. 55–76, 1984.

[27] N. Tandon and A. Chaudhary, "A Review of Vibration and Acoustic Measurement Methods for the Detection of Defects in Rolling Element Bearings," *Tribol. Int.*, vol. 32, no. 8, pp. 469–480, 1999.

[28] X. Zhang and J. Zhou, "Multi-Fault Diagnosis for Rolling Element Bearings Based on Ensemble Empirical Mode Decomposition and Optimized Support Vector Machines," *Mech. Syst. Signal Process.*, vol. 41, no. 1-2, pp. 127–140, 2013.

[29] G. Georgoulas, T. Loutas, C. D. Stylios, and V. Kostopoulos, "Bearing Fault Detection Based on Hybrid Ensemble Detector and Empirical Mode Decomposition," *Mech. Syst. Signal Process.*, vol. 41, no. 1-2, pp. 510–525, 2013.

[30] J. Zheng, J. Cheng, and Y. Yang, "Generalized Empirical Mode Decomposition and Its Applications to Rolling Element Bearing Fault Diagnosis," *Mech. Syst. Signal Process.*, vol. 40, no. 1, pp. 136–153, 2013.

[31] Z. Feng, M. J. Zuo, R. Hao, F. Chu, and J. Lee, "Ensemble Empirical Mode Decomposition-Based Teager Energy Spectrum for Bearing Fault Diagnosis," *J. Vib. Acoust.*, vol. 135, no. 3, pp. 031013-1–031013-21, 2013.

[32] W. Guo and P. W. Tse, "A Novel Signal Compression Method Based on Optimal Ensemble Empirical Mode Decomposition for Bearing Vibration Signals," *J. Sound Vib.*, vol. 332, no. 2, pp. 423–441, 2013.

[33] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A Review on Empirical Mode Decomposition in Fault Diagnosis of Rotating Machinery," *Mech. Syst. Signal Process.*, vol. 35, no. 1-2, pp. 108–126, 2013.

[34] M. Amarnath and I. R. P. Krishna, "Empirical Mode Decomposition of Acoustic Signals for Diagnosis of Faults in Gears and Rolling Element Bearings," *IET Sci. Meas. Technol.*, vol. 6, no. 4, pp. 279–287, 2012.

[35] G. F. Bin, J. J. Gao, X. J. Li, and B. S. Dhillon, "Early Fault Diagnosis of Rotating Machinery Based on Wavelet Packets-Empirical Mode Decomposition Feature Extraction and Neural Network," *Mech. Syst. Signal Process.*, vol. 27, pp. 696–711, 2012.

[36] W. Guo, P. W. Tse, and A. Djordjevich, "Faulty Bearing Signal Recovery from Large Noise Using a Hybrid Method Based on Spectral Kurtosis and Ensemble Empirical Mode Decomposition," *Measurement*, vol. 45, no. 5, pp. 1308–1322, 2012.

[37] Y. Lei, Z. He, and Y. Zi, "EEMD method and WNN for Fault Diagnosis of Locomotive Roller Bearings," *Expert Syst. Appl.*, vol. 38, no. 6, pp. 7334–7341, 2011.

[38] M. Zvokelj, S. Zupan, I. Prebil, "Multivariate and Multiscale Monitoring of Large-Size Low-Speed Bearings Using Ensemble Empirical Mode Decomposition Method Combined with Principal Component Analysis," *Mech. Syst. Signal Process.*, vol. 24, no. 4, pp. 1049–1067, 2010.

[39] P. K. Kankar, S. C. Sharma, and S. P. Harsha, "Fault Diagnosis of Rolling Element Bearing Using Cyclic Autocorrelation and Wavelet Transform," *Neurocomputing*, vol. 110, pp. 9–17, 2013.

[40] P. Konar and P. Chattopadhyay, "Bearing Fault Detection of Induction Motor Using Wavelet and Support Vector Machines (SVMs)," *Appl. Soft Comput.*, vol. 11, no. 6, pp. 4203–4211, 2011.

[41] K. Feng, Z. Jiang, W. He, and Q. Qin, "Rolling Element Bearing Fault Detection Based on Optimal Antisymmetric Real Laplace Wavelet," *Measurement*, vol. 44, no. 9, pp. 1582–1591, 2011.

[42] P. K. Kankar, S. C. Sharma, and S. P. Harsha, "Rolling Element Bearing Fault Diagnosis Using Wavelet Transform," *Neurocomputing*, vol. 74, pp. 1638–1645, 2011.

[43] P. K. Kankar, S. C. Sharma, and S. P. Harsha, "Rolling Element Bearing Fault Diagnosis Using Autocorrelation and Continuous Wavelet Transform," *J. Vib. Control*, vol. 17, no. 14, pp. 2081–2094, 2011.

[44] J. Rafiee, M. A. Rafiee, and P. W. Tse, "Application of Mother Wavelet Functions for Automatic Gear and Bearing Fault Diagnosis," *Expert Syst. Appl.*, vol. 37, no. 6, pp. 4568–4579, 2010.

[45] W. Su, F. Wang, H. Zhu, Z. Zhang, and Z. Guo, "Rolling Element Bearing Faults Diagnosis Based on Optimal Morlet Wavelet Filter and Autocorrelation Enhancement," *Mech. Syst. Signal Process.*, vol. 24, no. 5, pp. 1458–1472, 2010.

[46] K. Teotrakool, M. J. Devaney, and L. Eren, "Adjustable-Speed Drive Bearing-Fault Detection via Wavelet Packet Decomposition," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 8, pp. 2747–2754, 2009.

[47] Y. Feng and F. S. Schlindwein, "Normalized Wavelet Packets Quantifiers for Condition Monitoring," *Mech. Syst. Signal Process.*, vol. 23, no. 3, pp. 712–723, 2009.

[48] R. Yan, R. X. Gao, and X. Chen, "Wavelets for Fault Diagnosis of Rotary Machines: A Review with Applications," *Signal Process.*, vol. 96, pp. 1–15, 2014.

[49] J. Yu, "Local and Nonlocal Preserving Projection for Bearing Defect Classification and Performance Assessment," *IEEE Trans. Ind. Electron.*, vol. 59, no. 5, pp. 2363–2376, 2012.

[50] L. Jiang, J. Xuan, and T. Shi, "Feature Extraction Based on Semi-Supervised Kernel Marginal Fisher Analysis and Its Application in Bearing Fault Diagnosis," *Mech. Syst. Signal Process.*, vol. 41, no. 1-2, pp. 113–126, 2013.

[51] S. Dong and T. Luo, "Bearing Degradation Process Prediction Based on the PCA and Optimized LS-SVM Model," *Measurement*, vol. 46, no. 9, pp. 3143–3152, 2013.

[52] G. Georgoulas, M. O. Mustafa, I. P. Tsoumas, J. A. Antonino-Daviu, V. Climente-Alarcon, C. D. Stylios, and G. Nikolakopoulos, "Principal Component Analysis of the Start-Up Transient and Hidden Markov Modeling for Broken Rotor Bar Fault Diagnosis in Asynchronous Machines," *Expert Syst. Appl.*, vol. 40, no. 17, pp. 7024–7033, 2013.

[53] K. Feng, Z. Jiang, W. He, and B. Ma, "A Recognition and Novelty Detection Approach Based on Curvelet Transform, Nonlinear PCA and SVM with Application to Indicator Diagram Diagnosis," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 12721–12729, 2011.

[54] M. Zvokelj, S. Zupan, and I. Prebil, "Non-linear Multivariate and Multiscale Monitoring and Signal Denoising Strategy Using Kernel Principal Component Analysis Combined with Ensemble Empirical Mode Decomposition Method," *Mech. Syst. Signal Process.*, vol. 25, no. 7, pp. 2631–2653, 2011.

[55] V. H. Nguyen and J. -C. Golinval, "Fault Detection Based on Kernel Principal Component Analysis," *Eng. Struct.*, vol. 32, no. 11, pp. 3683–3691, 2010.

[56] M. Zhao, X. Jin, Z. Zhang, and B. Li, "Fault Diagnosis of Rolling Element Bearings via Discriminative Subspace Learning: Visualization and Classification," *Expert Syst. Appl.*, vol. 41, no. 7, pp. 3391–3401, 2014.

[57] Z. -B. Zhu and Z. -H. Song, "A Novel Fault Diagnosis System Using Pattern Classification on Kernel FDA Subspace," *Expert Syst. Appl.*, vol. 38, no. 6, pp. 6895–6905, 2011.

[58] D. G. Ece and M. Basaran, "Condition Monitoring of Speed Controlled Induction Motors Using Wavelet Packets and Discriminant Analysis," *Expert Syst. Appl.*, vol. 38, no. 7, pp. 8079–8086, 2011.

[59] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z. -H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg, "Top 10 Algorithms in Data Mining," *Knoewl. Inf. Syst.*, vol. 14, pp. 1–37, 2008.

[60] V. Kosse and A. C. C. Tan, "Development of Test Facilities for Verification of Machine Condition Monitoring Methods for Low Speed Machinery," in *Proc. World Cong. Eng. Asset Manag.*, July 11-14 2006, Gold Coast, Australia, pp. 192-197

[61] Y. -H. Kim, A. C. C. Tan, J. Mathew, V. Kosse, and B. S. Yang, "A Comparative Study on the Application of Acoustic Emission Technique and Acceleration Measurements for Low Speed Condition Monitoring," in *Proc. Asia-Pacific Vib. Conf.*, Aug. 8 2007, Sapporo, Japan.

[62] E. Y. Kim, A. C. C. Tan, B. S. Yang, and V. Kosse, "Experimental Study on Condition Monitoring of Low Speed Bearings: Time Domain Analysis," in Proc. Australasian Cong. Appl. Mechatronics., Dec. 10-12 2007, Brisbane, Australia.

[63] L. -S. Law, J. H. Kim, W. Y. H. Liew, and S. -K. Lee, "An Approach Based on Wavelet Packet Decomposition and Hilbert-Huang Transform (WPD-HHT) for Spindle Bearings Condition Monitoring," *Mech. Syst. Signal Process.*, vol. 33, pp. 197–211, 2012.

[64] X. Yao, L. Herrera, S. Ji, K. Zou, and J. Wang, "Characteristic Study and Time-Domain Discrete-Wavelet-Transform Based Hybrid Detection of Series DC Arc Faults," *IEEE Trans. Power Electron.*, vol. 29, no. 6, pp. 3103–3115, 2014.

[65] J. Manikandan and B. Venkataramani, "Evaluation of Multiclass Support Vector Machine Classifiers Using Optimum Threshold-Based Pruning Technique," *IET Signal Process.*, vol. 5, no. 5, pp. 506–513, 2011.

[66] I. Aydin, M. Karakose, and E. Akin, "A Multi-Objective Artificial Immune Algorithm for Parameter Optimization in Support Vector Machine," *Appl. Soft Comput.*, vol. 11, no. 1, pp. 120–129, 2011.

[67] I. Steinwart and C. Scovel, "Mercer's Theorem on General Domains: On the Interaction between Measures, Kernels, and RKHSs," *Constr. Approx.*, vol. 35, no. 3, pp. 363–417, 2012.

[68] J. D. Rodriguez, A. Perez, and J. A. Lozano, "Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 3, pp. 569–575, 2010.

Myeongsu Kang received BS and MS degrees in computer engineering and information technology in 2008 and 2010, respectively, from the University of Ulsan in Ulsan, South Korea, where he is a currently PhD student of electrical, electronics, and computer engineering. His current research interests include machinery fault diagnosis and condition monitoring, high-performance multimedia signal processing, and application-specific SoC design.

Jaeyoung Kim received a BS in electrical, electronics, and computer engineering in 2012 from the University of Ulsan, South Korea, where he is currently a MS student of electrical, electronics, and computer engineering. His current interests include expert and decision support systems, and high-performance multimedia signal processing.

Jong-Myon Kim (M'98) received a BS in electrical engineering from Myongji University in Yongin, Korea, in 1995, and MS in electrical and computer engineering from the University of Florida in Gainesville in 2000, and a PhD in electrical and computer engineering from the Georgia Institute of Technology in Atlanta in 2005. He is an Associate Professor of IT convergence at the University of Ulsan, Korea. His research interests include multimedia specific processor architecture, fault diagnosis and condition monitoring, parallel processing, and embedded systems. He is a member of IEEE and the IES Society.

Andy C. C. Tan received his BSc (Eng) and PhD degrees in mechanical engineering from the University of Westminster, London. His research interests include noise and vibration condition monitoring and sensors for machine and structural health monitoring. He applied adaptive signal processing and the blind deconvolution algorithms to enhance the desired signals corrupted by noise into incipient fault detection and machine diagnostics/prognostics. These algorithms together with acoustic emission sensors are currently being used in low speed machinery condition monitoring and bridge structures health monitoring. He is expanding his research in vibration control into developing bi-ventricular assist device as an artificial heart to enhance the life style and prolong live for final stage heart failure patients. His recent interest is the study of electrical impendence properties of Carbon-nano tube (CNT) for applications as sensors for bridge structure health monitoring and measurement of dynamics properties of artificial heart pump. He is a professor of mechanical engineering at the Faculty of Science and Engineering of the Queensland University of Technology and his academic interests include dynamics of mechanical systems, noise and vibrations and mechanism design. He is a Fellow of the Institution of Engineers, Australia.

Eric Y. Kim received the BS, MS, and PhD degrees in mechanical engineering from Pukyong National University, Korea, in 1997, 1999, and 2004, respectively. Through 10 years of academic researches and 3 years industry practical experience, his expertise extends to machine diagnostics/prognostics, signal processing, rotor-dynamics, condition monitoring and reliability engineering. He is currently working as a reliability engineer at CMOS Northparkes copper and gold mine in Australia.

Byeong-Keun Choi is a Professor at the Department of Energy Mechanical Engineering at Gyeongsang National University in Korea. He received his PhD degree in Mechanical Engineering from Pukyong National University, Korea, in 1999. Dr. Choi worked at Arizona State University as an Academic Professional from 1999 to 2002. Dr. Choi's research interests include vibration analysis and optimum design of rotating machinery, machine diagnosis and prognosis and acoustic emission.