Reliable Fault Diagnosis for Incipient Low-speed Bearings Using Fault Feature Analysis Based on a Binary Bat Algorithm

Myeongsu Kang[1*], Jaeyoung Kim[2*], and Jong-Myon Kim[3*], Andy C.C. Tan[4**], Eric Y. Kim[5***], Byeong-Keun Choi[6****]

Affiliation:

[*]School of Electrical, Electronics, and Computer Engineering, University of Ulsan, Ulsan, South Korea

[**]Cooperative Research Centre (CRC) for Integrated Engineering Asset Management, Queensland University of Technology, Australia
[***]Asset Management Department, CMOC Northparkes Mine, Australia

[****]Department of Energy and Mechanical Engineering, Gyeongsang National University, Tongyeong City, South Korea

Address:

[1,2]Bldg. #7, Room#304-1, 93 Daehak-ro, Nam-gu, Ulsan 680-749, South Korea

[3]Bldg. #7, Room #308, 93 Daehak-ro, Nam-gu, Ulsan 680-749, South Korea
[4]Northparkes Mine, PO Box 995, Parkes, NSW 2870, Australia
[5]Queensland University of Technology, 2 George St. Brisbane, QLD 4001, Australia
[6]38 Cheondaegukchi-gil, Tongyeong-si, Gyeongsangnam-do 650-160, South Korea

Email:

[1]ilmareboy@gmail.com

[2]kjy7097@gmail.com

[3]jmkim07@ulsan.ac.kr
[4]a.tan@qut.edu.au
[5]yonghan.kim@gmail.com

[6]bgchoi@gnu.ac.kr

Corresponding author:

Jong-Myon Kim, School of Electrical Engineering, University of Ulsan, Ulsan, South Korea

Email: jmkim07@ulsan.ac.kr

Tel: 82-52-259-2217

Fax: 82-52-259-1687

Abstract

This paper proposes a highly reliable fault diagnosis scheme for incipient low-speed rolling element bearing failures, which consists of fault feature calculation, discriminative fault feature analysis, and fault classification. The proposed approach first computes wavelet-based fault features including the relative wavelet packet node energy and wavelet packet node entropy by applying wavelet packet transform to an incoming acoustic emission (AE) signal. The proposed approach then sorts out the most discriminative fault features from the originally produced feature vector by using discriminative fault feature analysis based on a binary bat algorithm (BBA). Finally, the proposed approach employs one-against-all multi-class support vector machines (OAA MCSVMs) to identify multiple low-speed rolling element bearing defects. This study compares the proposed BBA-based dimensionality reduction scheme with four other dimensionality reduction methodologies in terms of classification performance. Experimental results show that the proposed methodology is superior to other dimensionality reduction approaches, yielding an average classification accuracy of 94.9%, 95.8%, and 98.4% under bearing rotational speeds at 20 revolutions-per-minute (RPM), 80 RPM, and 140 RPM, respectively.

Keywords

Acoustic emission, binary bat algorithm, dimensionality reduction, incipient low-speed bearing fault diagnosis, multi-class support vector machines, wavelet packet transform

1. Introduction

Bearings have played a significant role in low-speed machines that are widely utilized in heavy industries, such as paper mills, steel pipe foundries, and wind-turbine power plants, in order to support heavy loads and provide stationary rotational speeds [59]. Thus, unanticipated bearing defects can lead to severe motor breakdown and significant economic losses. To address this issue, reliable fault diagnosis for incipient bearing failures is important. Vibration signal analysis has been widely utilized for bearing fault diagnosis because it provides the most intrinsic information about diverse bearing defects [3, 4, 7, 26, 39, 48, 50, 53]. Likewise, current signature analysis has been an alternative for condition monitoring of bearings, offering two advantages [18, 23, 35, 46, 68, 69]: 1) high sensitivity to diverse mechanical failures with its non-intrusive monitoring ability, and 2) low-cost failure diagnosis, because it needs no special devices to be installed on the motor. Although these analysis methods have shown satisfactory performance for diagnosing diverse bearing failures, they focused only on identifying bearing faults under high rotational speeds, which are hundreds to thousands of revolutions-per-minute (RPM). This is because these methods have problems capturing useful descriptions about low-speed bearing failures from very feeble vibration and current signals. To address this problem, acoustic emission (AE) has been attractive for fault diagnosis in low-speed bearings, because AE is highly useful for capturing low-energy signals [6, 10–13, 44, 47, 57, 59, 60, 63]. More specifically, according to Tandon et al. [57] and Yoshioka et al. [63], AE can capture intrinsic symptoms of diverse bearing defects before they appear on the bearing's surface. Hence, in this study we exploit AE signals for early identification of incipient defects in low-speed rolling element bearings.

During the past few decades, signal processing-based fault diagnosis methodologies have been popular studies for identifying bearing failures, such as a crack or spall on raceways of a bearing. The following three steps are crucial for bearing fault diagnosis: fault feature calculation, discriminative fault feature analysis, and fault classification. In the fault feature computation phase, calculation of statistical parameters including mean, standard deviation, kurtosis, and skewness is performed in order to describe diverse symptoms of bearing defects by exploring time domain analysis [26, 44, 47, 48, 59, 60], frequency domain analysis [4, 7, 10, 24, 54, 59, 60], and time-frequency domain analysis [1, 5, 6, 14, 17,

19, 21, 22, 27–29, 33, 35, 37–39, 49, 53, 56, 58, 65, 67, 71]. Due to the fact that signals (e.g., vibration, current, AE) acquired for bearing defect diagnosis have non-stationary property, time-frequency domain analysis has been recently of increasing interest in research on revealing the most informative description about bearing failures. Hence, fault feature extraction has been performed via the time-frequency domain analyses including wavelet transform (WT) and empirical mode decomposition (EMD). Specifically, wavelet-based fault features have been extensively utilized for failure diagnosis [14, 16, 27–29, 33, 49, 56, 58, 62].

In general, the dimensionality of a feature vector including these wavelet-based fault features is high. However, this high-dimensional feature vector may have unuseful fault features which can increase the number of misclassifications among different bearing failures, and thus dimensionality reduction of the feature vector is required while keeping the most useful information about diverse bearing failures. To address this issue, several methods such as principal component analysis (PCA) [8, 15, 20, 26, 45, 59, 72] and linear discriminant analysis (LDA) [9, 26, 59, 66, 70] have been introduced to reduce the dimensionality of the feature vector. PCA, which is one of the unsupervised analysis methods, is good for dimensionality reduction, and the resultant principal components via PCA can be alternatives to optimal fault features for early diagnosis of bearing defects. However, PCA has a shortcoming when it comes to preserving discriminative properties of bearings, because it lacks an inter-class separability estimation process. On the other hand, LDA, which is one of the supervised analysis techniques, can preserve discriminative information by exploiting within-class and between-class scatter matrices. As a result, it generally offers better classification results for fault diagnosis than those obtained by the PCA-family approaches, such as PCA and kernel PCA [26, 66]. In addition, Jin et al. [26] and Zhao et al. [66] presented trace ratio linear discriminant analysis (TR-LDA) which is an orthogonal variant of LDA and eliminates redundant information from the scatter matrices in LDA. While the TR-LDA yielded satisfactory performance for identifying various bearing defects, it has a shortcoming for analyzing fault features not having Gaussian distribution. In practice, non-Gaussian fault features are often observed in many machinery diagnosis problems. Accordingly, the TR-LDA has been extended to effectively analyze non-Gaussian fault features. The extended TR-

LDA exploits two new scatter matrices characterizing the inter-class separability and the intra-class compactness, which are generated from intrinsic and penalty graphs. However, the extended TR-LDA often has a problem characterizing the inter-class separability as the penalty graph cannot reflect the neighborhood relationships among various classes, and this property inherent in the extended TR-LDA can be a major cause for degrading classification performance. To overcome these shortcomings in conventional dimensionality reduction approaches, this paper proposes a discriminative fault feature analysis based on a binary bat algorithm (BBA), which cooperates well with one-against-all multi-class support vector machines (OAA MCSVMs), where the SVM is a binary classifier using labeled information. The proposed discriminative fault feature analysis identifies an optimal set of fault features in an initially produced feature vector for diagnosing low-speed bearing failures. Finally, this proposed methodology discriminates various bearing defects by using an optimal set of fault features as an input of OAA MCSVMs, which are generally effective for achieving higher classification performance with limited training data [61].

The rest of this paper is organized as follows. Section 2 introduces diverse bearing defects and illustrates a test rig for experiments. Section 3 presents the proposed fault diagnosis methodology using fault feature analysis based on a BBA, and Section 4 evaluates its effectiveness in terms of classification performance, and compares the classification accuracy of the proposed approach with that of other state-of-the-art methods. Finally, Section 5 concludes this paper.

2. Diverse Rolling Element Bearing Defects and Experimental Setup

To identify incipient rolling element bearing defects, data obtained from a low-speed machinery fault simulator that enables modeling of bearing faults under different load conditions at different rotational speeds developed by *CRC-IEAM, Queensland University of Technology (QUT)* was used in the tests, as shown in Fig. 1(a) [34]. The test rig allows a range of bearing and gear faults to be simulated at low speeds with loads applied radially and by hydraulic devices. The test rig can record both vibration and Acoustic Emission (AE) signals, with the sensors mounted on top of the bearing housing, as shown in Fig. 1(b). In this paper, only AE data obtained using a wideband AE sensor

(Physical Acoustics Corp PAC R3a sensor) with a frequency range of 25-530 KHz was used. Fig. 1(c) illustrates a data acquisition system which is capable of 18-bit, 10-MHz analog-to-digital conversion, where a laptop was connected to a PCI board. To diagnose multiple bearing defects, various seeded bearing defects developed by QUT by using a diamond bit and a grinder with an air-grinding tool were used in this paper [31, 32]. Fig. 2 shows cracks and spalls on either raceways or the roller of cylindrical roller bearings (i.e., SKF NF307). In addition, a defect-free bearing was utilized as a reference case in this study. In total, this paper used six different types of AE signals under different load conditions (i.e., 500-N and 2-kN) at different rotational speeds (20, 80, and 140 RPM) and 90 records of 1.5-second AE signals sampled at 500 kHz for each case.

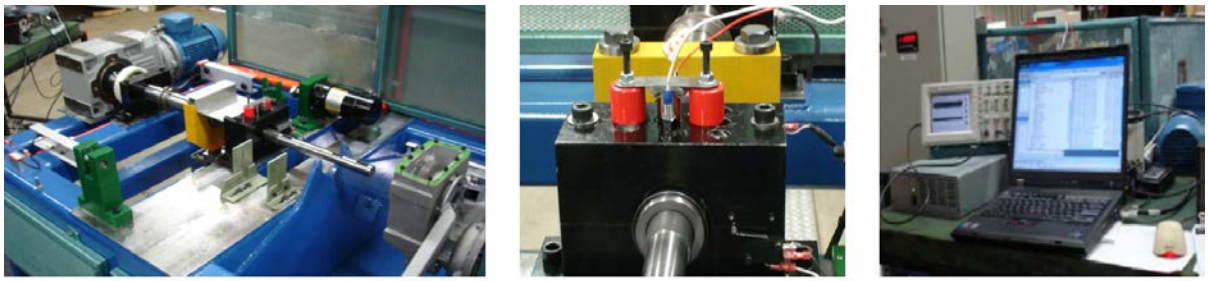

Fig. 1. (a) A low-speed machinery fault simulator developed by QUT [31, 32], (b) an AE sensor to record continuous AE signals, and (c) the data acquisition system used in this study.

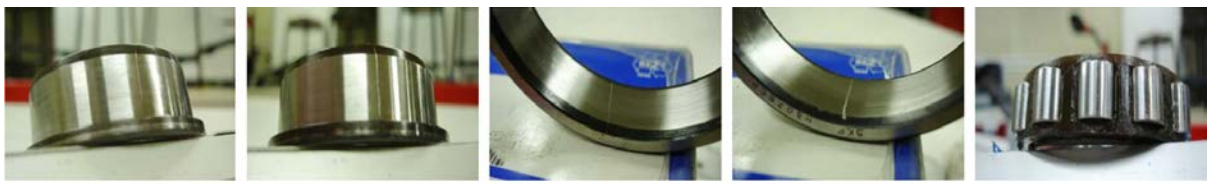

Fig. 2. Various seeded bearing failures [31, 32]. (a) Hair-line crack on inner raceway (CIR, 0.1 mm), (b) small-line spall on inner raceway (SIR, 0.6 mm), (c) hair-line crack on outer raceway (COR, 0.1 mm), (d) small-line spall on outer raceway (SOR, 0.7 mm), and (e) medium-line spall on roller (MSR, 1.6 mm).

In real industry, we can often observe overhung machines whose shafts are generally supported by two bearings (e.g., spherical or cylindrical roller bearings), and the bearings are one of the most

frequently failed components due to high vibration caused by mass unbalance and excessive loads. Specifically, the two bearings are differently loaded in overhung machines: a slight load is applied to the drive-end bearing while a relatively heavy load is applied to the non-drive-end bearing with approximately 3 times as much load as the drive-end bearing. Hence, it is significant to conduct research on diagnosing various bearing failures under different load conditions in order to offer a guideline for recognizing which bearing (i.e., the drive-end bearing or the non-drive-end bearing) is defective and what kinds of bearing failures are detected at either the drive-end bearing or the non-drive-end bearing in overhung machines. In this study, we totally identify 10 different bearing defects and two defect-free bearings: six (i.e., five bearing failures and a defect-free bearing) under a 500-N load and another six (i.e., five bearing failures and a defect-free bearing) under a 2-kN load.

3. Proposed Fault Diagnosis Methodology for Rolling Element Bearings

As mentioned in Section 1, the proposed fault diagnosis methodology includes fault feature extraction, fault feature analysis based on a BBA, and fault classification using OAA MCSVMs. This section details each step in the proposed scheme.

3.1. Fault Feature Extraction

To effectively capture AE signals for diagnosing various bearing failures, an AE sensor should be placed on a non-rotating element in machinery, such as the bearing housing. Although the bearing housing is the nearest non-rotating element when acquiring the signals for diagnosis, signal attenuation occurs due to its distance from the source of bearing defects. This signal attenuation calls a signal decomposition tool to explore intrinsic symptoms of various bearing defects in mid- and high-frequency regions. Discrete wavelet transform (DWT) is effective for analyzing defective information in AE signals owing to its decomposition ability, splitting signals into different frequency regions. In particular, wavelet packet transform (WPT) decomposes signals into uniform frequency bands, and thus it is more effective for achieving intrinsic information of bearing failures in mid- and high-

frequency bands than DWT [36]. Thus, we utilize WPT for the purpose of fault feature extraction, such as the relative wavelet packet node energy (RWPNE) and wavelet packet node entropy (WPNE). Both RWPNE and WPNE are used to effectively reveal disorder behaviors in signals, which are useful for representing symptoms of bearing failures [62]. In this study, we first perform three-level WPT to decompose a 1.5-second AE signal, resulting in eight terminal nodes. As previously mentioned, since the most useful information revealing various bearing defects mostly exists in mid- and high-frequency bands, we exploit the last six terminal nodes in order to calculate fault features such as RWPNE and WPNE. When analyzing the signals with wavelet transform-based decomposition tools, we should carefully choose a mother wavelet function because it greatly influences the analysis result. In this study, Daubechies 20 (or db20), which is one of Daubechies wavelet family, is used to decompose signals because it well matches to the recorded AE signals for diagnosis, as depicted in Fig. 3.

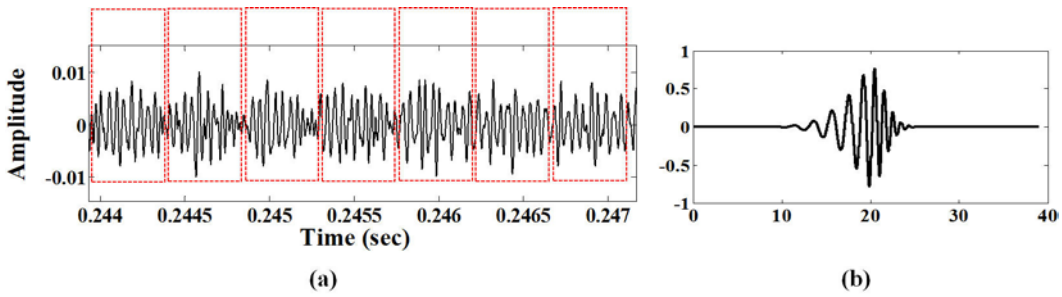

Fig. 3. (a) Waveform of an AE signal and (b) a Daubechies 20-tap wavelet (db20).

As fault features for identifying various bearing failures, this study first computes RWPNEs as follows:

$$RWPNE(j) = \sum_{i=1}^{N_{coeffs}} w_{i,j}^2 \Big/ \sum_{j=1}^{N_{tnodes}} \sum_{i=1}^{N_{coeffs}} w_{i,j}^2, \quad (1)$$

where N_{coeffs} is the number of wavelet coefficients, N_{tnodes} is the number of terminal nodes considered for computing fault features in this study, and $w_{i,j}$ is the ith wavelet coefficient in the jth terminal

node.

Then, this study calculates WPNEs as follows:

$$WPNE(j) = -\sum_{i=1}^{N_{coeffs}} p_j(i) \log_2 p_j(i), \tag{2}$$

where $p_j(i) = w_{i,j}^2 \Big/ \sum_{i=1}^{N_{coeffs}} w_{i,j}^2$. Because this study concentrates only on mid- and high-frequency bands which correspond to the last six terminal nodes, 12 fault features including six RWPNEs and six WPNEs are used to construct an initial feature vector for identifying various failures of bearings.

3.2. Dimensionality Reduction based on a Binary Bat Algorithm

For discriminative fault feature analysis, the proposed fault diagnosis approach employs a BBA [43], a heuristic method using the echolocation behavior of bats. The proposed BBA-based approach selects the most discriminative fault features to achieve the highest fault classification performance for low-speed rolling element bearings, where the BBA can identify an optimal feature vector for highly accurate low-speed bearing failure diagnosis by using the position vector of a bat, where each element of the position vector corresponds to the absence or presence of fault features (or statistical parameters) describing symptoms of rolling bearing defects. To select discriminative fault features, the BBA efficiently represents the position vector by using a set of binary values and an n_f-bit encoded position vector (i.e., 0 or 1), which is used to construct an optimal feature vector, where n_f is the total number of fault features. In this study, n_f is set to 12 because the proposed approach utilizes six RWPNEs and six WPNEs as fault features. For instance, if the value of the i^{th} element of the position vector is 1, we consider the i^{th} statistical parameter of fault features as one of the most discriminative fault features for identifying bearing failures, and vice versa.

In general, the BBA works with a set of initially generated position vectors (or solutions) called a *population*, and generates the best solution after a series of iterative computations. Fig. 4 presents the overall process of the BBA for analyzing fault features, consisting of the following five steps.

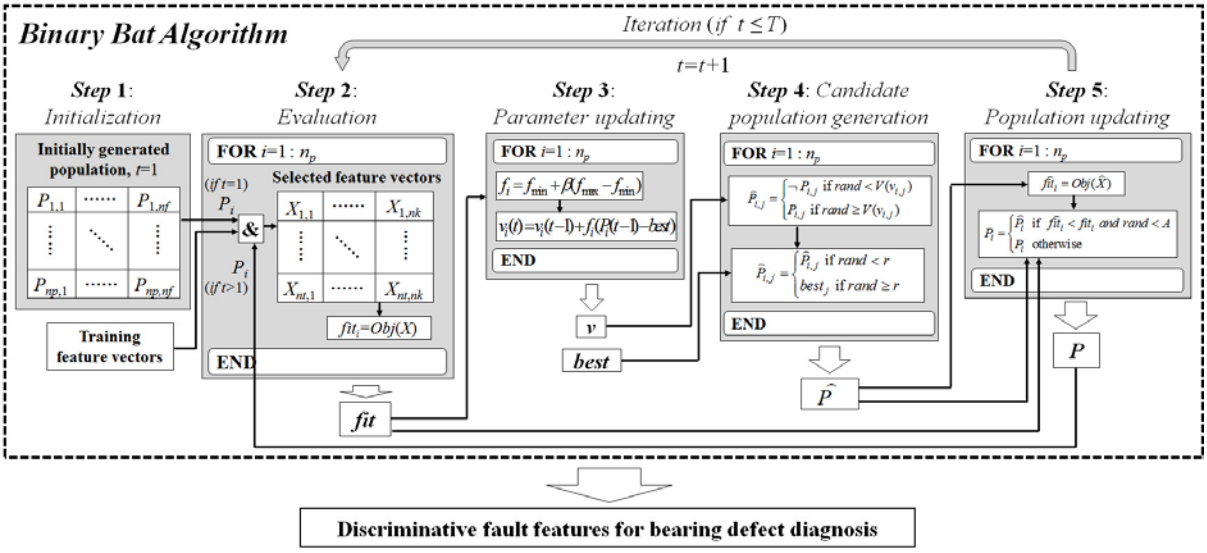

Fig. 4. The overall process of the binary bat algorithm for analyzing fault features, which sorts out the most discriminative fault features in a given feature vector for identifying incipient low-speed bearing defects.

■ *Step* 1: *Initialization*

The BBA sets an initial population, P, by randomly generating a certain number of position vectors (i.e., one position vector for each bat). The size of P is $n_p \times n_f$, where n_p is the number of position vectors (or the number of bats). As depicted in Fig. 4, the initial population P is the result of *Step* 1, and is used for all the remaining steps. Moreover, the BBA is based on the aforementioned echolocation behavior of bats, which is idealized by describing their behavior as they find prey. In general, a bat randomly flies with an arbitrary velocity at its position with a varying frequency, and it searches for prey by changing its velocity and frequency. Consequently, it is necessary to initialize the frequency and velocity of each bat in the population.

■ *Step* 2: *Evaluation*

In *Step* 2, the i^{th} position vector (or the position vector of bat i), P_i, is evaluated in the current population, where i=1, 2, …, n_p. An operator (&) in Fig. 4 is used to take a subset of fault features in the given feature vector by the i^{th} position vector, and an $n_t \times n_k$ matrix X is constructed that consists of fault features resulting from taking the & operation between the i^{th} position vector and each n_f-dimensional feature vector in a training dataset, in which n_t is the total number of training samples and n_k is the number of selected fault features by the i^{th} position vector (or the number of 1's in the

i^{th} position vector). Finally, an objective function evaluates the quality (or fitness) of the i^{th} position vector using the resultant matrix *X*. As shown in Fig. 4, *fit* is an n_p-dimensional fitness vector indicating the quality of position vectors in the current population, and it is further utilized in the next steps. Section 3.3 provides more details about designing a suitable objective function.

■ *Step* 3: *Parameter updating*

As mentioned in *Step* 1, the BBA iteratively updates frequency and velocity information to idealize the bat's echolocation behavior as follows [41]:

$$f_i = f_{\min} + \beta\left(f_{\max} - f_{\min}\right), \tag{3}$$

$$v_i\left(t\right) = v_i\left(t-1\right) + f_i\left\{P_i\left(t-1\right) - best\right\}, \tag{4}$$

where f_i and $v_i(t)$ are a random frequency and the velocity of bat *i* at iteration *t*, respectively. Likewise, $P_i(t\text{-}1)$ is the position vector of bat *i* at iteration *t*-1, and β is a random value within the interval [0, 1]. In addition, *best* is the position vector of the bat that yields the minimum fitness value. In summary, *Step* 3 outputs the n_p-dimensional velocity vector *v* (see *Step* 3 in Fig. 4), including updated velocities of each bat and the best solution in the current population (or *best* in *Step* 3 in Fig. 4). These two vectors are utilized to generate a candidate position vector for each bat in *Step* 4.

■ *Step* 4: *Candidate population generation*

In the BBA, each artificial bat modifies its own position in order to find prey. To model this in the BBA, it is necessary to flip bits of the position vector. However, the BBA cannot perform position updating for bats by adding the updated velocities to the previous positions, since it can only change positions from 0 to 1 or vice versa. Hence, several researchers have investigated ways to change the positions of bats with the probability of their velocities by using a transfer function, which maps velocity values to probability values [30, 42, 51]. According to Mirjalili et al. [41], large absolute values of bat velocities mean that bats need to fly far away to find prey and to switch their positions. Thus, a transfer function should offer a high probability of changing the position for a large absolute value of the velocity, and vice versa. To mathematically formulate this, Mirjalili et al. proposed a v-shaped transfer function as follows:

$$V(x) = \left| 2/\pi \cdot \tan^{-1}\left(\pi/2 \cdot x\right) \right|. \tag{5}$$

The position of bat *i* is then modified, which is expressed as follows:

$$\hat{P}_{i,j} = \begin{cases} \neg P_{i,j}, & if \quad rand < V\left(v_{i,j}\right) \\ P_{i,j}, & if \quad rand \geq V\left(v_{i,j}\right) \end{cases}, \tag{6}$$

$$\hat{P}_{i,j} = \begin{cases} \hat{P}_{i,j}, & if \quad rand < r \\ best_j, & if \quad rand \geq r \end{cases}, \tag{7}$$

where $\hat{P}_{i,j}$ is the j^{th} element of the position vector of bat *i*, where j=1, 2, …, n_k, $v_{i,j}$ is the j^{th} element of the velocity vector of bat *i*, *rand* is a random value ranging from 0 to 1, and *r* is a fixed constant within the interval [0, 1]. In this study, we set *r* to 0.9 for a divergent search for solutions. Finally, an $n_p \times n_k$ candidate position matrix is generated in *Step* 4. Then, *Step* 5 determines whether each element of the matrix should be switched or not, resulting in the final modified position matrix at iteration *t*.

■ *Step* 5: *Population updating*

This step updates the population for the next iteration by accepting or rejecting each element of the previously generated candidate position matrix. Specifically, a new population including modified position vectors for the bats is generated with the following rule:

$$P_i = \begin{cases} \hat{P}_i, & if \quad \hat{fit}_i < fit_i \ and \ rand < A \\ P_i, & otherwise \end{cases}, \tag{8}$$

where $\hat{fit}_i = Obj\left(\hat{X}\right)$, and *A* is any constant value ranging from 0 to 1, which controls the acceptability of the modified position vectors. In this study, we define *A* as 0.9 in order to more frequently reflect the modified position vectors for bats, rather than preserve the previously generated position vectors in *Step* 2. Similar to the matrix *X* in *Step* 2, $\hat{X}$ is a new $n_t \times n_k$ matrix consisting of fault features resulting from taking the & operation between the i^{th} candidate position vector produced in *Step* 4 and each n_f-dimensional feature vector in the training dataset. Consequently, $\hat{fit}_i$ is a fitness value to evaluate fault features selected by the candidate position vector of bat *i*. Finally, *Steps* 2 to 5 are repeated iteratively until the maximum iteration limit *T* is

reached.

To precisely evaluate the solution by the BBA, a proper design of the objective function is needed, and in turn this study investigates a methodology to appropriately design the objective function by exploiting the property of the SVM. As mentioned in Section 1, the SVM is a binary classifier to separate two different classes by finding a hyperplane with the largest margin between them in high-dimensional feature space. Based on this hyperplane, test samples are classified into one of the two classes. Let $s_j \in R^d$ and $l_j \in \{-1,+1\}$, $\forall j$=1, 2, …, n be a set of d-dimensional real valued training samples and labels corresponding to each training sample, respectively, where n is the number of training samples. To search for the optimal hyperplane when discriminating two classes, it is necessary to solve the following minimization optimal problem [2]:

$$\arg\min_{w,\xi_j}\left\{\frac{1}{2}w^T w + C\sum_{j=1}^{n}\xi_j\right\}, \tag{9}$$

$$\textit{subject to } l_j\left(w^T\phi\left(s_j\right)+b\right)\geq 1-\xi_j, \quad \xi_j \geq 0, \ \forall j=1,2,...,n,$$

where w is a vector to the hyperplane, ϕ is a function that maps the original feature space into the high-dimensional nonlinear feature space, b is a constant variable such that $b/\|w\|$ indicates the Euclidean distance from the origin of the high-dimensional nonlinear feature space to the hyperplane. Likewise, the slack variable ξ_j controls the training errors and the penalty variable C tunes the generalization capability. By applying Lagrange optimization to (9), the minimization optimal problem is written as follows [2]:

$$\arg\max_{\alpha_j}\left\{\sum_{j=1}^{n}\alpha_j - \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}\alpha_j\alpha_k l_j l_k \phi\left(s_j\right)^T\phi\left(s_k\right)\right\}, \tag{10}$$

$$\textit{subject to } \sum_{j=1}^{n}\alpha_j l_j = 0, \ 0\leq\alpha_j\leq C, \ \forall j=1,2,...,n,$$

where the α_j's are Lagrange multipliers, and s_j and s_k are two different samples in the training dataset. According to Mercer's theorem [55], a positive semi-definite kernel can replace with $\phi\left(s_j\right)\cdot\phi\left(s_k\right)$ in (10), where $\cdot$ means the inner product between two vectors:

$$k\left(s_j, s_k\right) = \phi\left(s_j\right) \cdot \phi\left(s_k\right). \tag{11}$$

The Gaussian radial basis function (RBF) kernel is one of the most widely used kernels with SVMs due to its satisfactory performance, and it is defined as follows:

$$k\left(s_j, s_k\right) = \exp\left(-\frac{1}{2\sigma^2}\left\|s_j - s_k\right\|^2\right), \tag{12}$$

where σ is a parameter to be carefully turned because of 1) losing non-linear power if σ is too small and 2) the lack of regularization if σ is too large.

Using the RBF kernel, the similarity can be measured between two input samples as expressed in (13) and (14):

$$k\left(a,b\right) \approx 1,\ \forall a,b \in C_j,\ \forall j = 1,2,...,L, \tag{13}$$

$$k\left(a,b\right) \approx 0,\ \forall a \in C_j,\ \forall b \in C_k,\ \forall j,k = 1,2,...,L,\ j \neq k, \tag{14}$$

where C_j is a set of samples in the class j, j=1, 2, …, L, where L is the number of classes. In this study, the results of (13) and (14) are treated as an inter-class RBF value and an intra-class RBF value, respectively. Figs. 5 and 6 show distribution of both intra-class RBF values and inter-class RBF values for two different sample sets, respectively. In the case where two classes are clearly separable as shown in Fig. 5(a), intra-class RBF values for each class are very high, while inter-class RBF values are close to 0.

On the other hand, both intra-class RBF values and inter-class RBF values are extensively distributed in the range from 0 to 1 when the two different classes are not clearly separable, as shown in Fig. 6(a). In this study, we utilize these two criteria to design the objective function. The first criterion, $RBF_{intraclass}\left(\hat{X}\right)$, is an average value of intra-class RBF values resulting from $\hat{X}$, where $\hat{X}$ is the resultant $n_t \times n_k$ matrix produced while the BBA analyzes fault features as mentioned in Section 3.2:

$$RBF_{intraclass}\left(\hat{X}\right) = \frac{1}{\sum_{j=1}^{L}\left|C_j\right|^2}\sum_{j=1}^{L}\sum_{k=1}^{|C_j|}\sum_{l=1}^{|C_k|} k\left(\hat{X}_{(j-1)\times|C_{j-1}|+k}, \hat{X}_{(j-1)\times|C_{j-1}|+l}\right), \tag{15}$$

where L is the number of classes and $\left|C_j\right|$ is the number of samples in class j (i.e., $n_t = L \times \left|C_j\right|$). The

second criterion, $RBF_{interclass}\left(\hat{X}\right)$, is an average value of inter-class RBF values resulting from $\hat{X}$, which is calculated by:

$$RBF_{interclass}\left(\hat{X}\right)=\frac{1}{\sum_{j=1}^{L}\sum_{\substack{k=1\\k\neq j}}^{L}\left|C_j\right|\left|C_k\right|}\sum_{j=1}^{L}\sum_{\substack{k=1\\k\neq j}}^{L}\sum_{l=1}^{|C_j|}\sum_{m=1}^{|C_k|}k\left(\hat{X}_{(j-1)\times|C_{j-1}|+l},\hat{X}_{(k-1)\times|C_{k-1}|+m}\right). \tag{16}$$

According to (13) and (14), if $RBF_{intraclass}\left(\hat{X}\right)$ is large and $RBF_{interclass}\left(\hat{X}\right)$ is small, both intra-class compactness and inter-class separability are improved. Thus, the objective function is designed with these two criteria in order to sort out the most discriminative fault features in an originally produced feature vector, and is defined as follows:

$$Obj\left(\hat{X}\right)=1-RBF_{intraclass}\left(\hat{X}\right)+RBF_{interclass}\left(\hat{X}\right). \tag{17}$$

The lowest value of the objective function corresponds to the minimal intra-class compactness and the maximal inter-class separability, and consequently this paper sorts out fault features in the given feature vector when the objective function yields the lowest value.

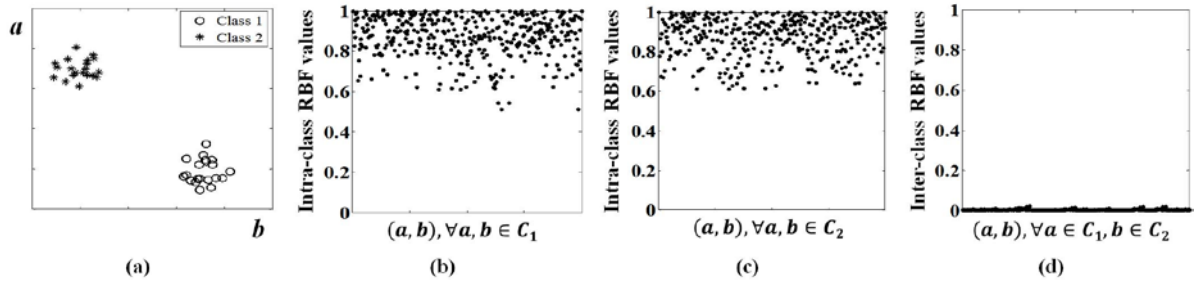

Fig. 5. Distribution of both intra-class RBF values and inter-class RBF values for two distinguishable classes.

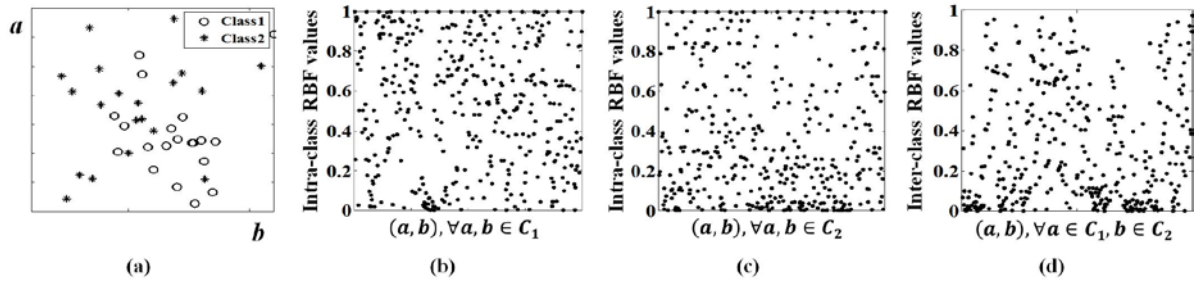

Fig. 6. Distribution of both intra-class RBF values and inter-class RBF values for two indistinguishable classes.

3.3. Classification

The SVM is a binary classifier to separate the test samples into one of two classes. Thus, the proposed approach utilizes multi-class SVMs (MCSVMs) to identify various (or multiple) bearing defects. In order to design MCSVMs, the following three approaches can be considered: one-against-one (OAO), one-against-all (OAA), and one-acyclic-graph (OAG) [40]. Among these approaches, we employ the OAA method because it is one of the most powerful techniques for multi-class classifiers, showing the high classification performance while reducing the testing time. In MCSVMs, SVM_j discriminates class j from the other classes, where j=1, 2, …, L and L is the number of classes as mentioned before, and all SVMs yield their own decision values during the classification process. In the OAA approach, the final decision is made by searing for an SVM structure that outputs the highest decision value.

4. Experimental Results

4.1. Training and Test Dataset Configuration

This study employs the k-fold cross validation (k-cv) scheme in order to estimate generalized classification accuracy [52], and the classification accuracy can be achieved by testing and training OAA MCSVMs using randomly divided k mutual folds from the initial data, denoted as D_1, D_2, …, D_k. In other words, a new fold is reserved for training OAA MCSVMs and the remaining folds are used to test OAA MCSVMs. This process is repeated k times to compute the final classification accuracy, which is the average value of the classification results committed in each fold (i.e., k = 3 in this study).

4.2. Classification Performance

For highly accurate bearing failure diagnosis, this paper uses discriminative fault feature analysis based on the BBA and validates its superiority by comparing with four conventional component analysis techniques, such as PCA, independent component analysis (ICA), kernel ICA (kICA), and

extended TR-LDA, in terms of the classification performance.

In general, an *n*-dimensional feature vector can be configured by using *n* components with the highest eigenvalues computed from a covariance matrix via these component analysis techniques. Although these approaches are effective for achieving satisfactory performance for bearing fault diagnosis, there is no general consensus as to the number of components that offers the maximum classification performance. This calls an exploration of the impact of discriminant (or principal or independent) components with regard to the classification performance, and finally this study builds an *n*-dimensional optimal feature vector with a set of *n* discriminant (or principal or independent) components providing the highest classification accuracy for the purpose of performance comparison. To do this, we randomly partition *k* mutual sub-folds from the fold that is reserved for training OAA MCSVMs in *k*-cv and explore the impact of varying numbers of discriminant (or principal or independent) components by training and testing OAA MCSVMs using these *k* mutual sub-folds, as shown in Fig. 7. For obtaining highly generalized classification accuracy, this study computes classification accuracy via 10 *k*-cv trials. Likewise, classification accuracy in this study is defined as follows:

$$Classification\ accuracy = \frac{\sum_{L} N_{truepositives}}{N_{samples}} \times 100\ (\%), \tag{18}$$

where L is the number of classes (i.e., L = 12 as mentioned in Section 2), $N_{truepositives}$ is the number of true positives, and $N_{samples}$ is the total number of samples for experiments (i.e., $N_{samples} = 1800 \times 12$). Moreover, the number of true positives indicates the number of bearing failures in class *j* that are correctly discriminated as class *j* in this study.

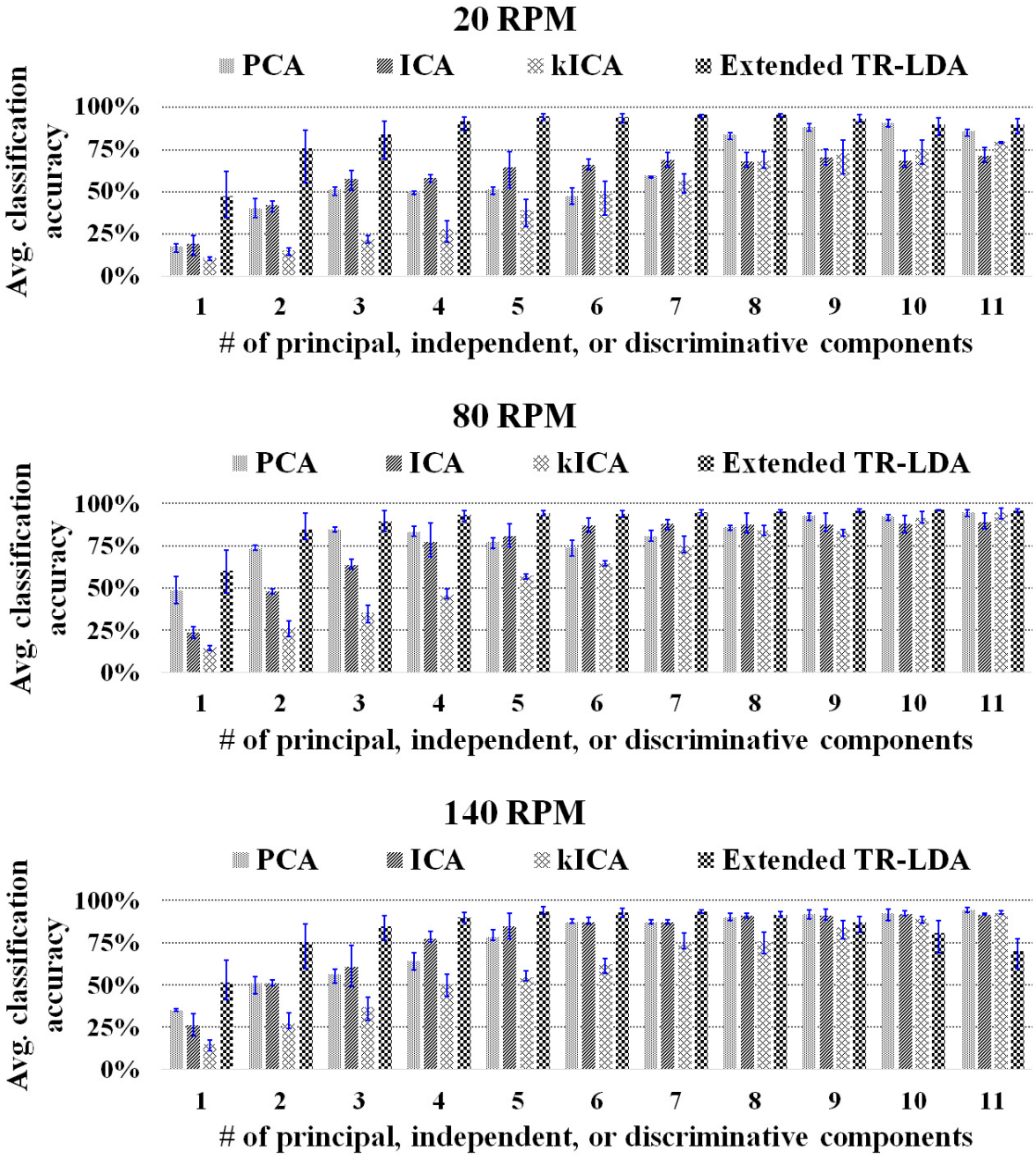

Fig. 7. Average classification accuracies with different numbers of principal, independent, or discriminant components under different bearing rotational speeds.

At each iteration in *k*-cv, this study decides optimal numbers of discriminant (or principal or independent) components from the ford that is reserved as the training set and computes the final classification accuracy by testing OAA MCSVMs with these optimal discriminant (or principal or independent) components for performance comparison, as depicted in Fig. 8. In addition, this study

selects optimal pairs of (C, σ) for OAA MCSVMs because these parameters can affect classification accuracy, as mentioned in Section 3.2. To address this issue, we employ a grid search algorithm that trains each SVM structure with a pair of (C, σ) in the cross-product of the following two sets and evaluates its performance: $C \in \left\{2^{-5}, 2^{-3}, 2^{-1}, 2, 2^{3}, 2^{5}, 2^{7}, 2^{9}, 2^{11}, 2^{13}, 2^{15}\right\}$ and $\sigma \in \left\{2^{-2}, 2^{-1}, 1, 2, 2^{2}, 2^{3}, 2^{4}, 2^{5}, 2^{6}, 2^{7}\right\}$. The grid search algorithm finally offers the best combination of (C, σ) for each SVM to obtain the maximum classification performance, and this study exploits the highest average classification accuracy of each approach for comparison. As we expected, both supervised (i.e., extended TR-LDA and the proposed approach) and unsupervised (i.e., PCA, ICA, and KICA) analysis methods decrease the average classification accuracy of bearing defects as rotational speed decreases, as shown in Fig. 8. This is highly correlated with both low-energy signals captured from bearings at low-speed and signal attenuation due to the distance between the AE sensor and the source of the bearing defects, as mentioned in Section 3.1.

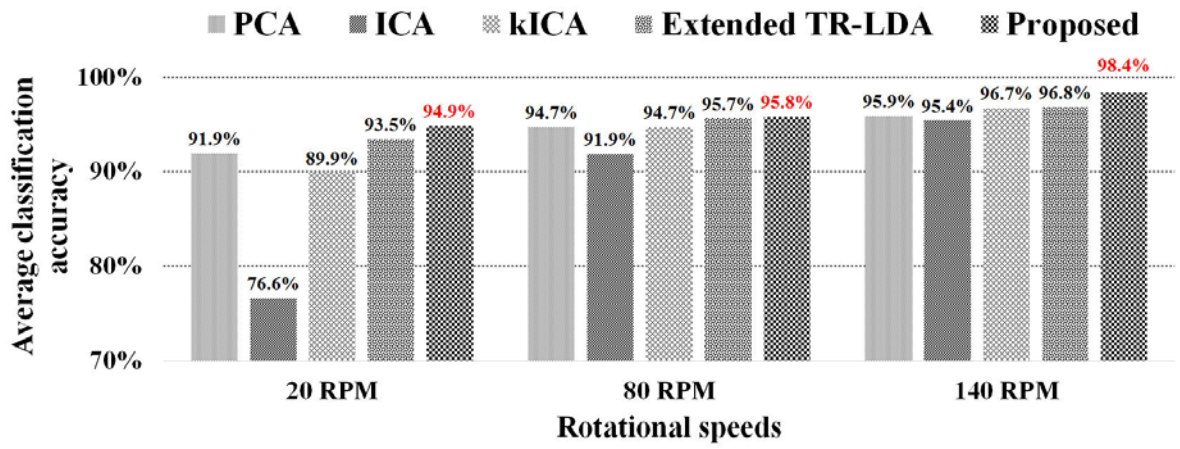

Fig. 8. Performance comparison between the proposed approach and other conventional methods with regard to average classification accuracy under various bearing rotational speeds.

In this study, we present confusion matrices in order to show detailed classification results (see Tables 1 to 5), and use sensitivity and specificity to indicate classification performance for each class of the proposed approach and other conventional approaches. Sensitivity (or true positive rate) and

specificity (or true negative rate) are defined as follows:

$$Sensitivity = \frac{N_{truepositives}}{N_{truenegatives} + N_{falsenegatives}} \times 100\ (\%). \tag{19}$$

$$Specificity = \frac{N_{truenegatives}}{N_{truenegatives} + N_{falsepositives}} \times 100\ (\%). \tag{20}$$

$N_{falsepositives}$, $N_{truenegatives}$, and $N_{falsenegatives}$ are the number of false positives, true negatives, and false negatives, respectively, and are defined as follows:

- $N_{truepositives}$: the number of bearing failures not in class *j* that are discriminated as class *j*,
- $N_{truenegatives}$: the number of bearing failures not in class *j* that are not discriminated as class *j*,
- $N_{falsenegatives}$: the number of bearing failures in class *j* that are not discriminated as class *j*.

Table 1
Classification results of the PCA-based bearing fault identification operating at 20 RPM.

	CIR1	SIR1	NB1	COR1	SOR1	MSR1	CIR2	SIR2	NB2	COR2	SOR2	MSR2
CIR1	1785	2	18	1	2	5	10	0	3	7	0	1
SIR1	0	1696	43	0	4	6	7	3	5	21	0	1
NB1	1	3	1321	1	5	305	4	0	415	3	10	11
COR1	0	9	18	1788	10	2	3	2	16	9	0	3
SOR1	0	0	5	0	1724	1	2	15	0	0	0	1
MSR1	0	0	299	0	0	1465	0	0	3	1	0	16
CIR2	14	13	0	0	1	0	1724	0	0	8	0	0
SIR2	0	0	0	2	18	0	0	1723	0	13	1	0
NB2	0	2	64	6	10	4	0	5	1345	2	0	3
COR2	0	67	0	0	0	0	37	50	0	1736	0	0
SOR2	0	8	31	2	26	1	13	2	13	0	1789	8
MSR2	0	0	1	0	0	11	0	0	0	0	0	1756
Sensitivity (%)	99.1	94.2	73.4	99.3	95.8	81.4	95.8	95.7	74.7	96.4	99.4	97.6
Specificity (%)	99.8	99.6	96.2	99.6	99.9	98.4	99.8	99.8	99.5	99.2	99.5	99.9
Accuracy (%)	91.9											

Table 2

Classification results of the ICA-based bearing fault identification operating at 20 RPM.

	CIR1	SIR1	NB1	COR1	SOR1	MSR1	CIR2	SIR2	NB2	COR2	SOR2	MSR2
CIR1	1637	7	9	0	5	2	18	0	2	39	0	2
SIR1	0	1243	115	47	46	59	169	61	37	238	0	15
NB1	1	20	1067	32	34	651	14	25	578	24	76	52
COR1	0	9	8	1556	0	0	0	4	9	38	0	4
SOR1	3	8	24	0	1545	2	16	76	8	3	74	0
MSR1	0	5	431	1	0	1023	0	1	198	10	11	72
CIR2	28	63	14	0	6	3	1410	0	0	18	0	0
SIR2	0	12	6	3	107	0	2	1500	8	65	0	0
NB2	0	5	48	2	6	2	2	14	943	1	0	4
COR2	130	416	7	157	32	0	156	108	0	1348	12	0
SOR2	0	11	53	2	19	5	13	12	15	14	1627	1
MSR2	1	1	18	0	0	53	0	0	2	2	0	1650
Sensitivity (%)	90.9	69.1	59.3	86.4	85.8	56.8	78.3	83.3	52.4	74.9	90.4	91.7
Specificity (%)	99.6	96.0	92.4	99.6	98.9	96.3	99.3	99.0	99.6	94.9	99.3	99.6
Accuracy (%)	76.6											

Table 3

Classification results of the kICA-based bearing fault identification operating at 20 RPM.

	CIR1	SIR1	NB1	COR1	SOR1	MSR1	CIR2	SIR2	NB2	COR2	SOR2	MSR2
CIR1	1745	4	6	0	10	1	11	2	4	3	1	8
SIR1	8	1184	10	2	15	4	57	15	5	50	0	3
NB1	0	0	1416	2	10	342	12	8	751	7	7	14
COR1	0	3	10	1433	5	3	0	7	7	10	0	4
SOR1	2	1	4	0	1522	2	13	73	2	1	24	2
MSR1	0	0	299	0	2	1437	3	5	1	2	3	55
CIR2	38	191	2	0	11	0	1566	0	0	28	0	0
SIR2	0	20	2	18	192	1	1	1633	2	22	0	0
NB2	0	7	41	12	6	3	3	4	1019	0	0	3
COR2	0	385	0	331	0	0	130	40	0	1673	0	1
SOR2	0	5	9	2	27	1	4	13	9	1	1764	2
MSR2	7	0	1	0	0	6	0	0	0	3	1	1708
Sensitivity (%)	96.9	65.8	78.7	79.6	84.6	79.8	87.0	90.7	56.6	92.9	98.0	94.9
Specificity (%)	99.8	99.2	94.2	99.8	99.4	98.1	98.6	98.7	99.6	95.5	99.6	99.9
Accuracy (%)	83.8											

Table 4

Classification results of the extended TR-LDA-based bearing fault identification operating at 20 RPM.

	CIR1	SIR1	NB1	COR1	SOR1	MSR1	CIR2	SIR2	NB2	COR2	SOR2	MSR2
CIR1	1783	2	7	1	1	2	5	0	1	3	0	1
SIR1	0	1720	20	0	7	4	41	4	9	14	0	3
NB1	1	2	1348	5	5	270	4	1	179	0	1	11
COR1	0	2	1	1769	3	0	0	0	2	5	1	0
SOR1	0	1	8	0	1734	0	1	12	1	0	3	1
MSR1	0	0	282	0	1	1491	1	0	0	1	2	13
CIR2	15	28	0	1	0	0	1709	1	2	30	0	0
SIR2	0	5	0	0	20	0	0	1758	0	9	0	0
NB2	0	6	77	10	8	9	0	1	1585	1	0	5
COR2	1	27	0	7	1	0	28	23	0	1737	0	0
SOR2	0	7	49	7	19	6	10	0	20	0	1790	8
MSR2	0	0	8	0	1	18	1	0	1	0	3	1758
Sensitivity (%)	99.1	95.6	74.9	98.3	96.3	82.8	94.9	97.7	88.1	96.5	99.4	97.7
Specificity (%)	99.9	99.5	97.6	99.9	99.9	98.5	99.6	99.8	99.4	99.6	99.4	99.8
Accuracy (%)	93.4											

Table 5

Classification results of the proposed BBA-based bearing fault identification operating at 20 RPM.

	CIR01	SIR1	NB1	COR1	SOR1	MSR1	CIR2	SIR2	NB2	COR2	SOR2	MSR2
CIR1	1800	0	2	0	1	0	1	0	0	0	0	2
SIR1	0	1784	0	0	0	0	0	0	0	0	0	0
NB1	0	0	1310	0	11	264	3	0	73	0	4	99
COR1	0	0	4	1793	2	0	2	0	0	0	0	0
SOR1	0	0	7	0	1764	0	5	0	0	0	3	1
MSR1	0	0	278	0	0	1487	0	0	3	0	0	12
CIR2	0	12	0	0	0	1	1773	0	0	14	0	0
SIR2	0	0	0	2	0	0	0	1800	0	0	0	0
NB2	0	0	58	0	0	1	0	0	1717	0	0	0
COR2	0	4	0	5	0	0	7	0	0	1786	0	0
SOR2	0	0	21	0	22	2	9	0	1	0	1792	3
MSR2	0	0	120	0	0	45	0	0	6	0	1	1683
Sensitivity (%)	100.0	99.1	72.8	99.6	98.0	82.6	98.5	100.0	95.4	99.2	99.6	93.5
Specificity (%)	100.0	100.0	97.7	100.0	99.9	98.9	99.9	100.0	99.7	99.9	99.7	99.1
Accuracy (%)	94.9											

As shown in Tables 1 to 5, all the fault feature analysis methodologies have difficulties analyzing extracted fault features of both NB1 (NB under a 500-N load) and MSR1 (MSR under a 500-N load) operating at 20 RPM, resulting in a number of misclassifications between NB1 and MSR1. To investigate performance degradation, this paper uses one-dimensional visualization for an input vector of OAA MCSVMs configured by both supervised approaches (i.e., PCA, ICA, and kICA) and unsupervised approaches (i.e., extended TR-LDA and the proposed approach). Analysis of the results is summarized as follows:

- Although it is possible to achieve higher classification accuracies by utilizing low-dimensional feature vectors for a few bearing defects in the PCA-based approach, rather than using high-dimensional feature vectors, this paper utilizes a 10-dimensional feature vector including 10 principal components for early identification of diverse bearing defects. A 10-dimensional feature vector offers the highest classification accuracy, which can be interpreted to mean that the PCA-based fault diagnosis method identifies as many bearing defects as possible with this feature vector (see Fig. 7). This is the primary reason why the PCA-based approach has difficulty in distinguishing NB1 from MSR1, and vice versa. Specifically, most fault features (i.e., from PC4 to PC10) are tightly overlapped between the two classes, as shown in Fig. 9. Thus, these features do not contribute much for identifying NB1 and MSR1. On the other hand, several PCs (i.e., PC1, PC2, and PC3) preserve discriminative properties of these two classes well, which are primary factors in achieving sensitivity of 73.4%.
- Similar to the PCA-based approach, the other methodologies also have problems classifying these two classes, since most of the fault features are not distinguishable, as shown in Fig. 9.

While other conventional component analysis approaches yield low sensitivities for distinguishing between NB2 and NB1, as shown in Tables 1 to 5, the proposed feature analysis methodology is highly effective, resulting in sensitivity of 95.4%. More specifically, the lack of measuring intra-class compactness and inter-class separability in both PCA and ICA-family approaches is the main reason why they cannot preserve discriminative properties of these two defect-free bearings, resulting in unclear and heavily overlapped features, as shown in Fig. 10. In contrast, extended TR-LDA reduces

the number of misclassifications from NB2 to NB1 by exploiting intra-class compactness and inter-class separability information. Although the extended TR-LDA achieves a satisfactory level of sensitivity for NB2, the proposed approach yields a much higher sensitivity than that of extended TR-LDA by extracting the most discriminative fault features in the given feature vector, as shown in Fig. 10.

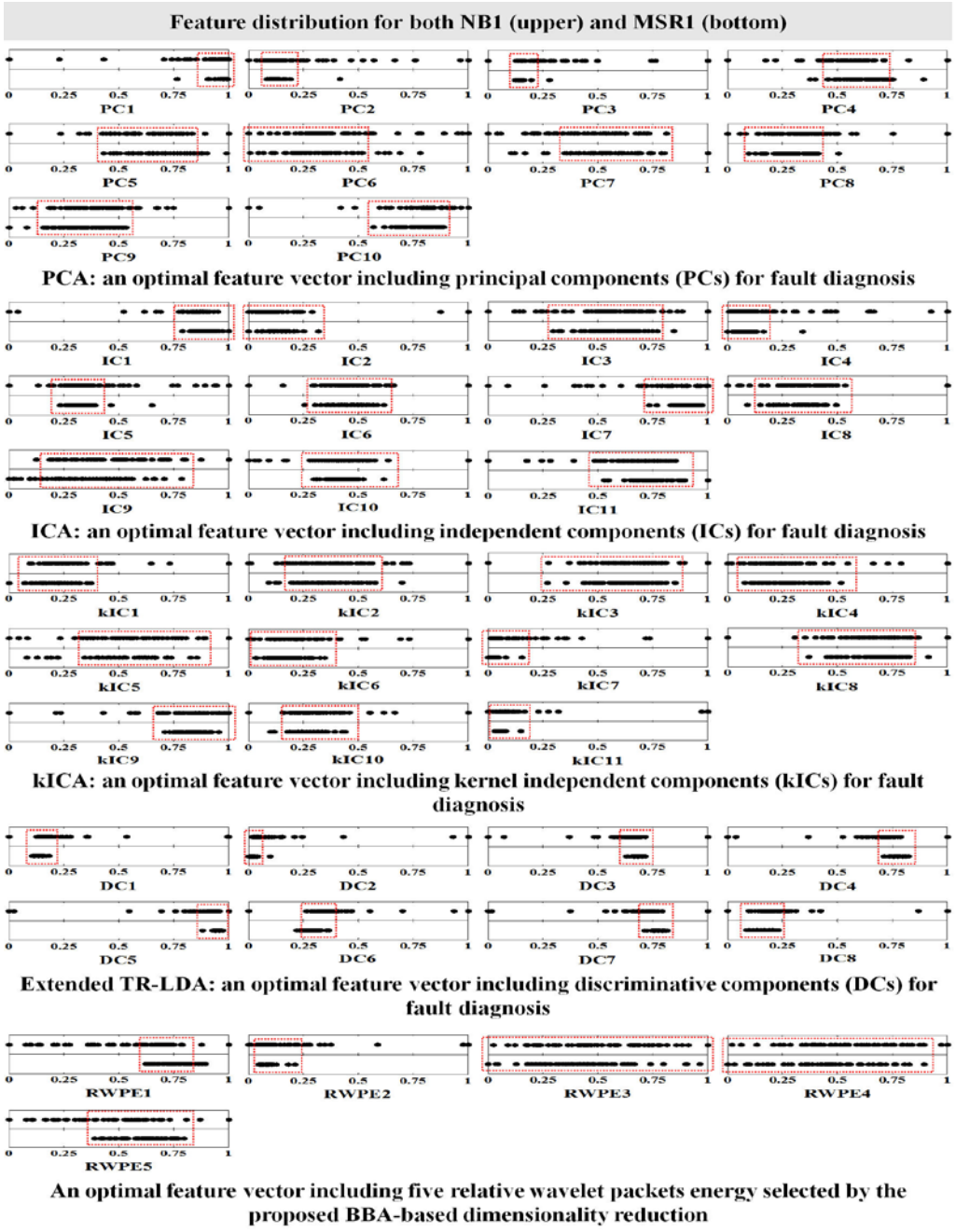

Fig. 9. One-dimensional visualization of inputs for MCSVMs configured by both the proposed BBA-based approach and other conventional methods. Overlapped features are shown in red dotted rectangles, which do not contribute to distinguishing between NB1 and MSR1, and vice versa. Likewise, input values are normalized from 0 to 1 by using min-max normalization in order to show feature distribution in the same range.

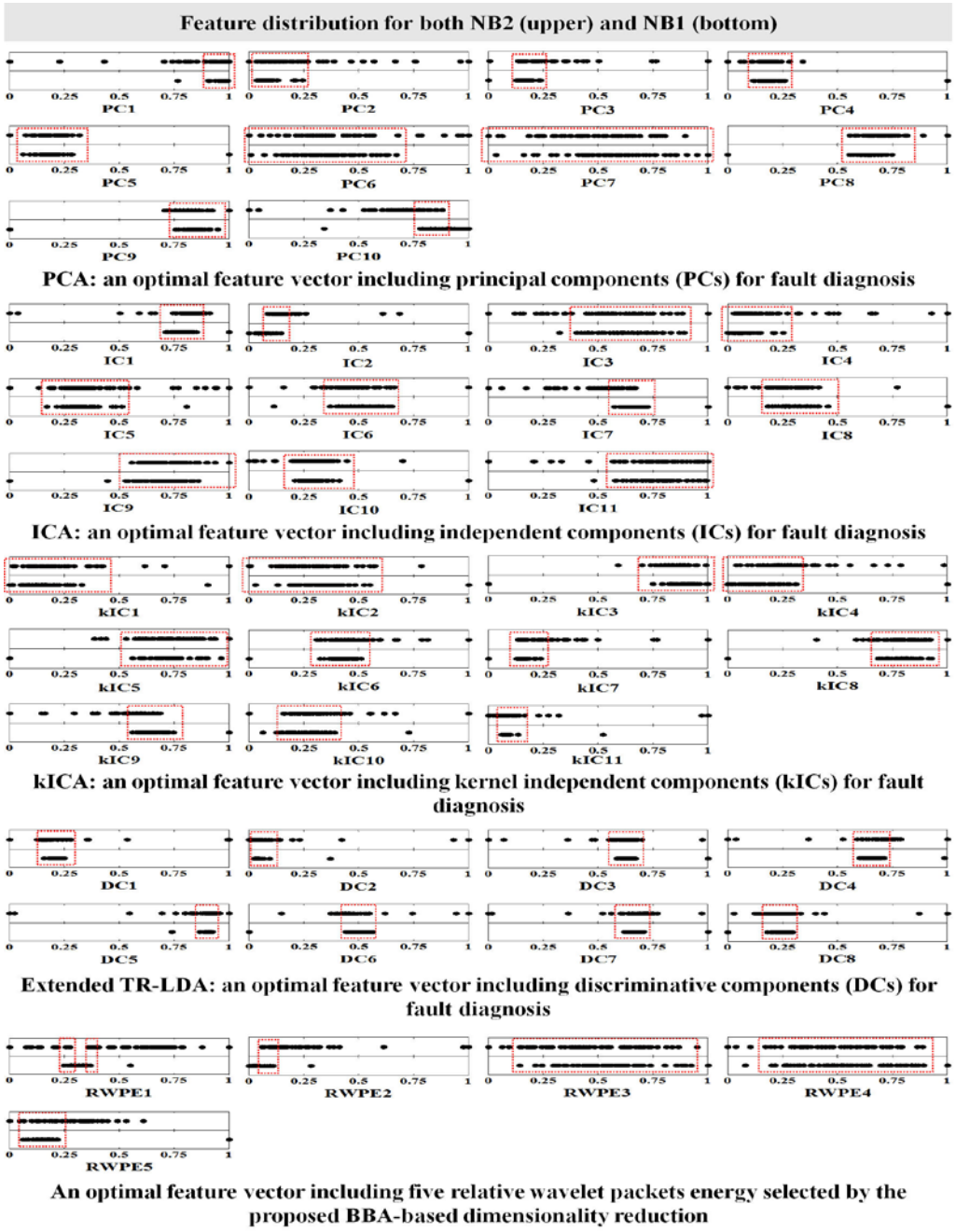

Fig. 10. One-dimensional visualization of inputs for MCSVMs to distinguish between NB2 and NB1.

5. Conclusions

This paper proposed a reliable fault diagnosis methodology for low-speed rolling element bearings, which is composed of fault feature extraction, BBA-based discriminative fault feature analysis, and fault classification. The proposed approach first computes wavelet-based fault features

(i.e., the relative wavelet packet node energy and wavelet packet node entropy) by applying WPT to AE signals. The proposed comprehensive diagnosis approach then sorts out the most discriminative fault features in the given feature vector in order to improve classification performance, and finally identifies multiple bearing defects by employing OAA MCSVMs. Experimental results indicated that the BBA-based feature analysis is superior to other conventional component analysis methods for the purpose of reliable bearing failure diagnosis, achieving an average classification accuracy of 94.9%, 95.8%, and 98.4% at rotational speeds of 20 RPM, 80 RPM, and 140 RPM, respectively.

Acknowledgements

This research was supported by the 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.

References

[1] M. Amarnath, I.R.P. Krishana, Empirical mode decomposition of acoustic signals for diagnosis of faults in gears and rolling element bearings, IET Science, Measurement and Technology, 6 (4) (2012) 279 – 287.

[2] I. Aydin, M. Karakose, E. Akin, A multi-objective artificial immune algorithm for parameter optimization in support vector machine, Applied Soft Computing, 11 (1) (2011) 120–129.

[3] I. Bediaga, X. Mendizabal, A. Arnaiz, J. Munoa, Ball bearing damage detection using traditional signal processing algorithms, IEEE Instrumentation and Measurement Magazine, 16 (2) (2013) 20–25.

[4] I. Bediaga, X. Mendizabal, I. Etxaniz, J. Munoa, An integrated system for machine tool spindle head ball bearing fault detection and diagnosis, IEEE Instrumentation and Measurement Magazine, 16 (2) (2013) 42–47.

[5] G.F. Bin, J.J. Gao, X.J. Li, B.S. Dhillon, Early fault diagnosis of rotating machinery based on wavelet packets – empirical mode decomposition feature extraction and neural network, Mechanical Systems and Signal Processing, 27 (2012) 696–711.

[6] W. Caesarendra, P.B. Kosasih, A.K. Tieu, B. –K. Choi, Condition monitoring of naturally damaged slow speed slewing bearing based on ensemble empirical mode decomposition, Journal of Mechanical Science and Technology, 27 (8) (2013) 2253–2262.

[7] M. Cococcioni, B. Lazzerini, S.L. Volpi, Robust diagnosis of rolling element bearings based on classification techniques, IEEE Transactions on Industrial Informatics, 9 (4) (2013) 2256–2263.

[8] S. Dong, T. Luo, Bearing degradation process prediction based on the PCA and optimized LS-SVM model, Measurement, 46 (9) (2013) 3143–3152.

[9] D.G. Ece, M. Basaran, Condition monitoring of speed controlled induction motors using wavelet packets and discriminant analysis, Expert Systems with Applications, 38 (7) (2011) 8079–8086.

[10] B. Eftekharnejad, M.R. Carrasco, B. Charnley, D. Mba, The application of spectral kurtosis on acoustic emission and vibrations from a defective bearing, Mechanical Systems and Signal Processing, 25 (1) (2011) 266–284.

[11] E. Elforjani, D. Mba, Condition monitoring of slow-speed shafts and bearings with acoustic emission, International Journal of Experimental Mechanics, 47 (2) (2011) 350–363.

[12] M. Elforjani, D. Mba, Accelerated natural fault diagnosis in slow speed bearings with acoustic emission, Engineering Fracture Mechanics, 77 (1) (2010) 112–127.

[13] M. Elforjani, D. Mba, Natural mechanical degradation measurements in slow speed bearings, Engineering Failure Analysis, 16 (1) (2009) 521–532.

[14] K. Feng, Z. Jiang, W. He, Q. Qin, Rolling element bearing fault detection based on optimal antisymmetric real Laplace wavelet, Measurement, 44 (9) (2011) 1582–1591.

[15] K. Feng, Z. Jiang, W. He, B. Ma, A recognition and novelty detection approach based on curvelet transform, nonlinear PCA and SVM with application to indicator diagram diagnosis, Expert Systems with Applications, 38 (10) (2011) 12721–12729.

[16] Y. Feng, F.S. Schlindwein, Normalized wavelet packets quantifiers for condition monitoring, Mechanical Systems and Signal Processing, 23 (3) (2009) 712–723.

[17] Z. Feng, M.J. Zuo, R. Hao, F. Chu, J. Lee, Ensemble empirical mode decomposition-based Teager energy spectrum for bearing fault diagnosis, Journal of Vibration and Acoustics, 125 (3) (2013) 031013-1–031013-21.

[18] L. Frosini, E. Bassi, Stator current and motor efficiency as indicators for different types of bearing faults in induction motors, IEEE Transactions Industrial Electronics, 57 (1) (2010) 244–251.

[19] G. Georgoulas, T. Loutas, C.D. Stylios, V. Kostopoulos, Bearing fault detection based on hybrid ensemble detector and empirical mode decomposition, Mechanical Systems and Signal Processing, 41 (1-2) (2013) 510–525.

[20] G. Georgoulas, M.O. Mustafa, I.P. Tsoumas, J.A. Antonino-daviu, V. Climente-alarcon, C.D. Stylios, G. Nikolakopoulos, Principal component analysis of the start-up transient and hidden Markov modeling for broken rotor bar fault diagnosis in asynchronous machines, Expert Systems with Applications, 40 (17) (2013) 7024–7033.

[21] W. Guo, P.W. Tse, A novel signal compression method based on optimal ensemble empirical mode decomposition for bearing vibration signals, Journal of Sound and Vibration, 332 (2) (2013) 423–441.

[22] W. Guo, P.W. Tse, A. Djordjevich, Faulty bearing signal recovery from large noise using a hybrid method based on spectral kurtosis and ensemble empirical mode decomposition, Measurement, 45 (5) (2012) 1308–1322.

[23] A. Ibrahim, M.E. Badaoui, F. Guillet, F. Bonnardot, A new bearing fault detection method in induction machines based on instantaneous power factor, IEEE Transactions. Industrial Electronics, 55 (12) (2008) 4252–4259.

[24] F. Immovilli, A. Bellini, R. Rubini, C. Tassoni, Diagnosis of bearing faults in induction machines by vibration or current signals: a critical comparison, IEEE Transactions on Industry Applications, 46 (4) (2010) 1350–1359.

[25] L. Jiang, J. Xuan, T. Shi, Feature extraction based on semi-supervised kernel marginal fisher analysis and its application in bearing fault diagnosis, Mechanical Systems and Signal Processing, 41 (1-2) (2013) 113–126.

[26] X. Jin, M. Zhao, T.W.S. Chow, M. Pecht, Motor bearing fault diagnosis using trace ratio linear discriminant analysis, IEEE Transactions on Industrial Electronics, 61 (6) (2014) 2441–2451.

[27] P.K. Kankar, S.C. Sharma, S.P. Harsha, Fault diagnosis of rolling element bearing using cyclic autocorrelation and wavelet transform, Neurocomputing, 110 (2013) 9–17.

[28] P.K. Kankar, S.C. Sharma, S.P. Harsha, Rolling element bearing fault diagnosis using wavelet transform, Neurocomputing, 74 (2011) 1638–1645.

[29] P.K. Kankar, S.C. Sharma, S.P. Harsha, Rolling element bearing fault diagnosis using autocorrelation and continuous wavelet transform, Journal of Vibration and Control, 17 (14) (2011) 2081–2094.

[30] J. Kennedy, R.C. Eberhart, A discrete binary version of the particle swarm algorithm, in: Proc. 1997 IEEE Int. Conf. Systems, Man, and Cybernetics, Orlando, USA, 1997, pp. 4104–4108.

[31] E.Y. Kim, A.C.C. Tan, J. Mathew, V. Kosse, B.S. Yang, A comparative study on the application of acoustic emission technique and acceleration measurements for low speed condition monitoring, in: Asia-Pacific Vibration Conference(APVC), Sapporo, Japan, 2007.

[32] E.Y. Kim, A.C.C. Tan, B.S. Yang, V. Kosse, Experimental study on condition monitoring of low speed bearings: time domain analysis, in: Australasian Congress on Applied Mechanics (ACAM), December, Brisbane, Australia, 2007.

[33] P. Konar, P. Chattopadhyay, Bearing fault detection of induction motor using wavelet and support vector machines (SVMs), Applied Soft Computing, 11 (6) (2011) 4203–4211.

[34] V. Kosse and A.C.C. Tan, Development of test facilities for verification of machine condition monitoring methods for low speed machinery, in: Int. Proc. of WCEAM, Gold Coast, Australia, 2006.

[35] E.C.C. Lau, H.W. Ngan, Detection of motor bearing outer raceway defect by wavelet packet transformed motor current signature analysis, IEEE Transactions on Instrumentation and Measurement, 59 (10) (2010) 2683–2690.

[36] L.–S. Law, J.H. Kim, W.Y.H. Liew, S.–K. Lee, An approach based on wavelet packet decomposition and Hilbert-Huang transform (WPD-HHT) for spindle bearings condition monitoring, Mechanical Systems and Signal Processing, 33 (2012) 197–211.

[37] Y. Lei, J. Lin, Z. He, M.J. Zuo, A review on empirical mode decomposition in fault diagnosis of rotating machinery, Mechanical Systems and Signal Processing, 35 (1-2) (2013) 108–126.

[38] Y. Lei, Z. He, Y. Zi, EEMD method and WNN for fault diagnosis of locomotive roller bearings, Expert Systems with Applications, 38 (6) (2011) 7334–7341.

[39] J. Liu, W. Wang, F. Golnaraghi, An extended wavelet spectrum for bearing fault diagnosis, IEEE Transactions on Instrumentation and Measurement, 57 (12) (2008) 2801–2812.

[40] J. Manikandan, B. Venkataramani, Evaluation of multiclass support vector machine classifiers using optimum threshold-based pruning technique, IET Signal Processing, 5 (5) (2011) 506–513.

[41] S. Mirjalili, S.M. Mirjalili, X.–S. Yang, Binary bat algorithm, Neural Computing and Applications (in press). doi:http://dx.doi.org/10.1007/s00521-013-1525-5.

[42] S. Mirjalili, S.Z.M. Hashim, BMOA: binary magnetic optimization algorithm, International Journal of Machine Learning and Computing, 2 (3) (2012) 204–208.

[43] R.Y.M. Nakamura, L.A.M. Pereira, K.A. Costa, D. Rodrigues, J.P. Papa, X.–S. Yang, BBA: a binary bat algorithm for feature selection, in Proc. IEEE Conf. Graphics, Patterns, and Images, Ouro Preto, 2012, pp. 291–297.

[44] S.A. Niknam, V. Songmene, Y.H.J. Au, The use of acoustic emission information to distinguish between dry and lubricated rolling element bearings in low-speed rotating machines, International Journal of Advanced Manufacturing Technology, 69 (9-12) (2013) 2679–2689.

[45] V.H. Nguyen, J.–C. Golinval, Fault detection based on kernel principal component analysis, Engineering Structures, 32 (11) (2010) 3683–3691.

[46] I.Y. Onel, M.E.H. Benbouzid, Induction motor bearing failure detection and diagnosis: park and Concordia transform approaches comparative study, IEEE/ASME Transactions on Mechatronics, 13 (2) (2008) 257–262.

[47] D.H. Pandya, S.H. Upadhyay, S.P. Harsha, Fault diagnosis of rolling element bearing with intrinsic mode function of acoustic emission data using APF-KNN, Expert Systems with Applications, 40 (10) (2013) 4137–4145.

[48] M.D. Prieto, G. Cirrincione, A.G. Espinosa, J.A. Ortega, H. Henao, Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks, IEEE Transactions on Industrial Electronics, 30 (8) (2013) 3398–3407.

[49] J. Rafiee, M.A. Fafiee, P.W. Tse, Application of mother wavelet functions for automatic gear and bearing fault diagnosis, Expert Systems with Applications, 37 (6) (2010) 4568–4579.

[50] A.S. Raj, N. Murali, Early classification of bearing faults using morphological operators and fuzzy inference, IEEE Transactions on Industrial Electronics, 60 (2) (2013) 567–574.

[51] E. Rashedi, H. Nezamabadi-pour, S. Saryzdi, BGSA: binary gravitational search algorithm, Natural Computing, 9 (3) (2010) 727–745.

[52] J.D. Rodriguez, A. Perez, J.A. Lozano, Sensitivity analysis of k-fold cross validation in prediction error estimation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 32 (3) (2010) 569–575.

[53] J. Seshadrinath, B. Singh, Investigation of vibration signatures for multiple fault diagnosis in variable frequency drives using complex wavelets, IEEE Transactions on Power Electronics, 29 (2) (2014) 936–945.

[54] A. Soualhi, G. Clerc, H. Razik, Detection and diagnosis of faults in induction motor using an improved artificial ant clustering technique, IEEE Transactions on Industrial Electronics, 60 (9) (2013) 4053–4062.

[55] I. Steinwart, C. Scovel, Mercer's theorem on general domains: on the interaction between measures, kernels, and RKHSs, Constructive Approximation, 35 (3) (2012) 363–417.

[56] W. Su, F. Wang, H. Zhu, Z. Zhang, Z. Guo, Rolling element bearing faults diagnosis based on optimal Morlet wavelet filter and autocorrelation enhancement, Mechanical Systems and Signal Processing, 24 (5) (2010) 1458–1472.

[57] N. Tandon, A. Chaudhary, A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings, Tribology International, 32 (8) (1999) 469–480.

[58] K. Teotrakool, M.J. Devaney, L. Eren, Adjustable-speed drive bearing-fault detection via wavelet packet decomposition, IEEE Transactions on Instrumentation and Measurement, 58 (8) (2009) 2747–2754.

[59] A. Widodo, B.–S. Yang, E.Y. Kim, A.C.C. Tan, J. Mathew, Fault diagnosis of low speed bearing based on acoustic emission signal and multi-class relevance vector machine, Nondestructive Testing and Evaluation, 24 (4) (2009) 313–328.

[60] A. Widodo, E.Y. Kim, J.–D. Son, B.–S. Yang, A.C.C. Tan, D.–S. Gu, B.–K. Choi, J. Mathew, Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine, Expert Systems with Applications, 36 (3) (2009) 7252–7261.

[61] 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, D. Steinberg, Top 10 algorithms in data mining, Knowledge and Information Systems, 14 (1) (2008) 1–37.

[62] R. Yan, R.X. Gao, X. Chen, Wavelets for fault diagnosis of rotary machines: a review with applications, Signal Processing, 96 (2014) 1–15.

[63] T. Yoshioka, T. Fujiwara, Application of acoustic emission technique to detection of rolling bearing failure, American Society of Mechanical Engineers, 14 (1984) 55–76.

[64] J. Yu, Local and nonlocal preserving projection for bearing defect classification and performance assessment, IEEE Transactions on Industrial Electronics, 59 (5) (2012) 2363–2376.

[65] X. Zhang, J. Zhou, Multi-fault diagnosis for rolling element bearings based on ensemble empirical mode decomposition and optimized support vector machines, Mechanical Systems and Signal Processing, 41 (1-2) (2013) 127–140.

[66] M. Zhao, X. Jin, Z. Zhang, B. Li, Fault diagnosis of rolling element bearings via discriminative subspace learning: visualization and classification, Expert Systems with Applications, 41 (7) (2014) 3391–3401.

[67] J. Zheng, J. Cheng, Y. Yang, Generalized empirical mode decomposition and its applications to rolling element bearing fault diagnosis, Mechanical Systems and Signal Processing, 40 (1) (2013) 136–153.

[68] W. Zhou, B. Lu, T.G. Habetler, R.G. Harley, Incipient bearing fault detection via motor stator current noise cancellation using wiener filter, IEEE Transactions on Industry Applications, 45 (4) (2009) 1309–1317.

[69] W. Zhou, T.G. Habetler, R.G. Harley, Bearing fault detection via stator current noise cancellation and statistical control, IEEE Transactions on Industrial Electronics, 55 (12) (2008) 4260–4269.

[70] Z.–B. Zhu, Z.–H. Song, A novel fault diagnosis system using pattern classification on kernel FDA subspace, Expert Systems with Applications, 38 (6) (2011) 6895–6905.

[71] 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, Mechanical Systems and Signal Processing, 24 (4) (2010) 1049–1067.

[72] M. Zvokelj, S. Zupan, I. Prebil, Non-linear multivariate and multiscale monitoring and signal denosing strategy using kernel principal component analysis combined with ensemble empirical mode decomposition method, Mechanical Systems and Signal Processing, 25 (7) (2011) 2631–2653.

Biographies:

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 a PhD student of electrical, electronic, and computer engineering. His current research interests include reliable fault diagnosis and condition monitoring of electrical machinery, high-performance multimedia signal processing, and application-specific SoC design.

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

Jong-Myon Kim received a BS in electrical engineering from Myongji University in Yongin, Korea, in 1995, an 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 Electrical Engineering 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.

Byeong-Keun Choi is a Professor at the Department of Energy Mechanical Engineering at Gyeongsang National University in Korea. He received his Ph.D. degrees 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.

Eric Y. Kim received the B.S., Master and Ph.D. 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 in CMOC Northparkes copper and gold mine in Australia.

Andy CC 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.

Table captions:

Table 2

Classification results of the PCA-based bearing fault identification operating at 20 RPM.

Table 2

Classification results of the ICA-based bearing fault identification operating at 20 RPM.

Table 3

Classification results of the kICA-based bearing fault identification operating at 20 RPM.

Table 4

Classification results of the extended TR-LDA-based bearing fault identification operating at 20 RPM.

Table 5

Classification results of the proposed BBA-based bearing fault identification operating at 20 RPM.

Figure captions:

Fig. 9. (a) A low-speed machinery fault simulator developed by QUT, (b) an AE sensor to record continuous AE signals, and (c) the data acquisition system used in this study.

Fig. 10. Various seeded bearing failures [31, 32]. (a) Hair-line crack on inner raceway (CIR, 0.1 mm), (b) small-line spall on inner raceway (SIR, 0.6 mm), (c) hair-line crack on outer raceway (COR, 0.1 mm), (d) small-line spall on outer raceway (SOR, 0.7 mm), and (e) medium-line spall on roller (MSR, 1.6 mm).

Fig. 11. (a) Waveform of an AE signal and (b) a Daubechies 20-tap wavelet (db20).

Fig. 12. The overall process of the binary bat algorithm for analyzing fault features, which sorts out the most discriminative fault features in a given feature vector for identifying incipient low-speed bearing defects.

Fig. 13. Distribution of both intra-class RBF values and inter-class RBF values for two distinguishable classes.

Fig. 14. Distribution of both intra-class RBF values and inter-class RBF values for two indistinguishable classes.

Fig. 15. Average classification accuracies with different numbers of principal, independent, or discriminant components under different bearing rotational speeds.

Fig. 16. Performance comparison between the proposed approach and other conventional methods with regard to average classification accuracy under various bearing rotational speeds.

Fig. 9. One-dimensional visualization of inputs for MCSVMs configured by both the proposed BBA-based approach and other conventional methods. Overlapped features are shown in red dotted rectangles, which do not contribute to distinguishing between NB1 and MSR1, and vice versa. Likewise, input values are normalized from 0 to 1 by using min-max normalization in order to show feature distribution in the same range.

Fig. 170. One-dimensional visualization of inputs for MCSVMs to distinguish between NB2 and NB1.