Discrete Wavelet Design for Bearing Fault Diagnosis Using Particle Swarm Optimization

Jalel Khelil[1], Khaled Khelil[2*], Messaoud Ramdani[1], Nadir Boutasseta[3]

[1] Department of Electronics, Badji-Mokhtar University, BP. 12, Annaba 23000, Algeria
[2] Faculty of Sciences and Technology, LEER Lab., Souk Ahras University, Souk Ahras 41000, Algeria
[3] Research Center in Industrial Technologies CRTI P.O. Box 64, Cheraga, Algiers 16000, Algeria

Corresponding Author Email: khaled.khelil@univ-soukahras.dz

ABSTRACT

Rolling bearings are widely used in a large variety of industrial applications. Therefore, it is necessary to provide an efficient fault detection and diagnosis mechanism to prevent component failure and poor performance during operation. This paper proposes a novel classification scheme based on the design of discrete wavelets best adapted to vibration signal analysis in order to identify and properly classify rolling bearing defects. Through polyphase representation of the wavelet filter bank, and using the particle swarm optimization (PSO) algorithm, the appropriate discrete wavelet associated filters are optimized to achieve the best fault classification accuracy. Simulation results show that the proposed wavelet design approach outperforms the well-known standard wavelets regardless the employed classifier and leads to an average fault classification improvement of about 2%.

1. INTRODUCTION

Industrial plants often contain rotating machinery including turbines, compressors, pumps, motors, etc. that should be permanently monitored to reduce possible faults and hence, improve their operational efficiency. Bearings, known for being among the most important parts of these rotating machines, are generally made up of an outer and inner ring, cage and rolling elements. These rolling bearings, whose failures are amongst the most common reasons for machines breakdown, often operate under extreme working conditions leading to defects and gradual decrease of performance [1].

Statistical studies have shown that many of the mechanical defects of rotating machinery are caused by rolling bearings, where inner and outer ring cracks are the most common faults representing about 90% of all the possible faults, while the rolling elements and cage cracking constitute nearly 10 % of all the defects [2]. To establish an efficient condition-based maintenance for bearing defects monitoring, various techniques such as vibration analysis, acoustic emission, temperature measurement, wear analysis, etc. have been proposed [3].

Generally, these techniques apply signal processing tools, including data acquisition, feature extraction, feature selection, and fault classification. Feature extraction step is a key factor affecting greatly the fault classification performance [4, 5]. Thus, different methods based on time domain, frequency domain, and time-frequency domain feature extractions have been suggested. However, time and frequency domain features have been shown to be relatively inappropriate for non-stationary signals which can be better processed using time-frequency domain techniques able to provide simultaneously, both time-domain and frequency-domain information [6]. To deal with non-stationary signals, including mechanical and bearing faulty signals, many time-frequency domains

techniques have been used, such as the short-time Fourier transform (STFT), the Wigner-Ville distribution (WVD), the empirical mode decomposition (EMD), and the wavelet transform (WT) [7-11]. More specifically, the WT based approach, due to its effectiveness, has been widely and successfully applied in monitoring mechanical diagnosis problems and bearing fault detection. Shan [12] adopted wavelet packet decomposition to extract energy features and used adaptive Fuzzy-SVM classifier to diagnose the faults of rolling bearings. Kumar et al. [13] examined the discrete wavelet transform (DWT) performance in diagnosing bearing defects using Neural Networks. Applying different wavelets such us Daubechies and Symlet wavelets (db8, db44, db4 and Sym10), they showed that de-noising the vibration raw signal using DWT, improves significantly the fault classification process. Mohammed et al. [14] verified the DWT effectiveness to quantify different non-stationary faulty signals and its ability to track the signal dynamics properly. To undertake the problem of bearing fault diagnosis using wavelets, Yan et al. [15] showed the advantages of the DWT with respect to the continuous wavelet transform (CWT) regarding computation time and early fault detection. A novel hybrid fault diagnostic method for roller bearing under multi-speed is presented using impulse modelling continuous wavelet transform (IMCWT) model [16]. Patil et al. [17] used DWT with ANN to detect and classify bearing faults where they employed Daubechies and Symlet wavelets (db4, db8, sym5 and sym8) to extract statistical features such as RMS, Skewness, Crest, and Kurtosis.

To the best of our knowledge, in all the reported research works relevant to fault detection and diagnosis (FDD), the issue of designing the wavelet filters for preprocessing the vibration signal, has not been examined. In this paper a new FDD approach, based on optimized wavelet filters using the PSO algorithm, is proposed. In other words, instead of using a

well-known standard wavelet such as Morlet, Daubechies, Symlet, or others for preprocessing the acquired vibration signals, we intend to design discrete wavelet filters that are the most suitable for rolling bearing diagnosis in terms of fault classification. Therefore, based on Mallat multiresolution analysis [18], this is accomplished by making use of Sherlock and Monro iterative algorithm [19] through the generation of two channel perfect reconstruction wavelet filter banks. The key idea of the procedure starts from the interesting Vaidyanathan factorization of the polyphase matrix [20], which consists of a parameterization of the space of the perfect reconstruction two-channel filter banks. Subsequently, a convenient reformulation [19] of such parameterization allows generating the 2M coefficients of the low-pass (LP) filter from M angular parameters θ_i taking values in the interval $[0, 2\pi[$ [21]. In this work, the optimal wavelet design is formulated as an optimization problem within the $\mathbb{R}^N$ space of the angular parameters θ_i. Hence, using the particle swarm optimization (PSO) algorithm, the discrete wavelet associated filters are optimized in terms of the bearing fault classification accuracy. Four Intelligent classifiers have been considered for the bearing defect classification viz., K Nearest neighbors (KNN), Fuzzy K-Nearest Neighbors (FKNN), Artificial Neural Networks (ANN) and Support Vector Machine (SVM).

The remaining of this paper is organized as follows: Section 2 provides a description of the experimental test rig employed in this work. Section 3 describes briefly the four employed classifiers. The discrete wavelet transforms principle and the polyphase representation of wavelet filter bank are presented in section 4, along with the Particle Swarm Optimization (PSO) algorithm and the wavelet filter design procedure. The obtained simulation results are discussed in section 5. Finally, the conclusions are drawn in section 6.

2. EXPERIMENTAL TEST RIG DESCRIPTION

In the present work, the experiments are carried out using the bearing test rig, shown in Figure 1, belonging to the Unit of Research in Advanced Materials (URMA) (Annaba, Algeria) of the Research Center in Industrial Technology (CRTI). Vibration signals are collected by considering five different states of the bearing: (i) healthy bearing; (ii) with inner race fault; (iii) with ball defect; (iv) with outer race fault; (v) combination of bearing elements defect.

Three ICP industrial (CTC AC 140-2D) accelerometers are employed, one placed on the top of a 0.37 kW three-phase induction motor, while the two others are mounted on the bearing housings to collect signals from both healthy and defective bearings.
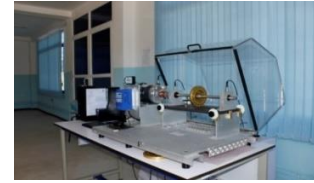

Figure 1. Vibration analysis test rig of the Unit of Research in Advanced Materials URMA/CRTI (Annaba, Algeria)

The characteristics of the used bearings are reported in Table 1.

Table 1. Bearing technical specifications

Bearing Model	ER12K-0114
Inner race diameter (mm)	29.31
Outer race diameter (mm)	47.00
Ball diameter (mm)	7.93
Contact angle (degree)	0°

The vibration signals are analyzed and stored using the Vibra Quest software (VQ-DT8) at a sampling frequency of 25.6 kHz and an acquisition time of 10 seconds for each operating condition. Figure 2 represents the time domain vibration signals for the five considered conditions of the employed bearings. Through a variable speed drive Lenze ESV371N01SXB, and considering the five bearing states, all tests are conducted at three different rotational speeds: 1800, 2100 and 2700 rpm, resulting in 15 temporal vibration signals of 262146 samples per signal. Therefore, for a purpose of classification, each signal is segmented into 50 non overlapping bins having 5240 samples each, giving rise to a database composed of 750 temporal vibration signals.
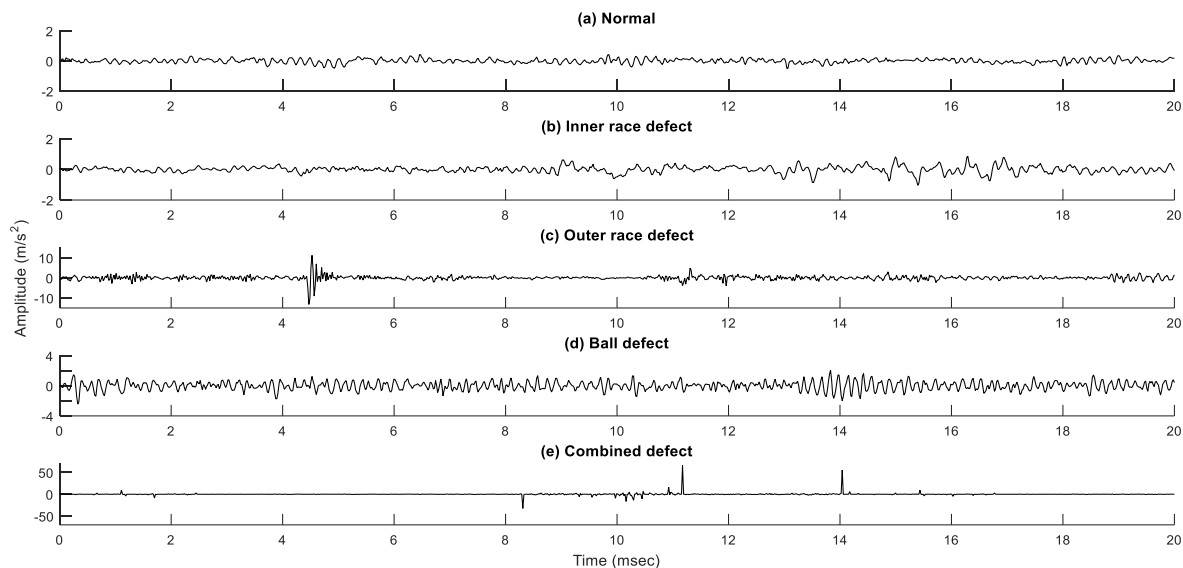
Figure 2. (a) Sampled vibration signal of normal bearing, (b) Inner race defect; (c) Outer race defect; (d) Ball defect; (e) Combined defect

3. BRIEF REVIEW OF KNN, FKNN, ANN AND SVM CLASSIFIERS

In the present study, to demonstrate the methodology proposed for bearing fault diagnosis and classification, four supervised machine learning techniques, i.e. KNN, FKNN, ANN and SVM are considered.

3.1 K-Nearest Neighbors (KNN)

The K-Nearest Neighbors (KNN) algorithm is one of the most important techniques for classification [22]. It's a simple non-parametric procedure for the assignment of a class label to the new input sample based on the class labels represented by the majority of the K nearest neighbors of the training samples in the feature space. The parameter K which specifies the number of nearest neighbors of the input sample can be optimized by various heuristic techniques. For the particular case when K is one, the input sample is simply assigned to the class of that single nearest neighbor. In this work, K=3 has been found to be the optimal value after testing several values from 1 to 10.

3.2 Fuzzy K-Nearest Neighbors (FKNN)

Compared to traditional classifiers which can only assign one class to the input sample, fuzzy classifiers assign a fuzzy vector that indicates the relative strength of each class. Based on the advantage of fuzzy set theory, an improved version of KNN method was proposed by Keller in 1985 and named as fuzzy KNN algorithm (FKNN) [23].

It provides not only the simplicity and practicability of classical KNN, but also uses the advantages of fuzzy logic theory. The fuzzy KNN algorithm allocates class membership to a sample vector instead of allocating the vector to a given class.

Let $\{x_1, \ldots, x_n\}$ be set of n sample vectors for which a fuzzy partition C defines the membership degree of each x_j in each of the C classes. Denoted by μ_{ij} for $i = 1, \ldots, C$ and $j = 1, \ldots, n$, it represents the degree of membership of the jth vector x_j in the ith class.

The following properties must hold for μ_{ij} to be a C partition:

$$\sum_{i=1}^{C} \mu_{ij} = 1$$
$$0 < \sum_{j=1}^{n} \mu_{ij} < n \quad (1)$$
$$\mu_{ij} \in [0,1]$$

In the classification step, for each new input sample x its K nearest neighbors are determined, and the assigned memberships $\mu_i(x)$ of the vector x are calculated according to:

$$\mu_i(x) = \frac{\sum_{j=1}^{K} \mu_{ij} \left(1/\|x - x_j\|^2 \right)}{\sum_{j=1}^{K} 1/\|x - x_j\|^2} \quad (2)$$

$\|x - x_j\|^2$ is the Euclidean distance between x and its jth nearest neighbor x_j. In this paper, the memberships μ_{ij} are determined using the constrained fuzzy membership [24]:

$$\mu_{ij} = \begin{cases} 0.51 + \left(\dfrac{n_j}{K}\right) * 0.49 & if\ i = j \\ \left(\dfrac{n_j}{K}\right) * 0.49 & if\ i \neq j \end{cases} \quad (3)$$

where, n_j represents the number of neighbors belonging to the jth class.

3.3 Artificial Neural Networks (ANN)

Artificial neural networks (ANN) are among the popular supervised machine learning techniques widely used in several problems with nonlinearities such as fault diagnosis and classification [25]. In this paper, a feedforward multilayer perceptron (MLP) composed of two hidden layers having 10 neurons each, is employed. The *NN* training is conducted using the Levenberg-Marquardt algorithm.

3.4 Support Vector Machine (SVM)

The Support vector machine (SVM) pioneered by Vapnik in the late 1960s is a powerful algorithm that has been successfully applied in several fields such as detection, recognition, prediction, machinery fault diagnosis, and so on [26]. SVM is a machine learning technique for small samples classification [25]. The SVM algorithm constructs an optimal separating hyperplane *f(x)* between data sets by resolving a constrained quadratic optimization problem based on the structural risk minimization (SRM) [27].

$$y = f(x) = W^T x + b = \sum_{i=1}^{N} W_i x_i + b \quad (4)$$

where, W is a *N*-dimensional vector and b is a scalar. The optimal separating hyperplane is tending to create the largest distance between the plane and the nearest data, that is, the maximum margin as illustrated in Figure 3. Through the conversion of the optimization problem with Kuhn-Tucker condition into the equivalent Lagrangian dual quadratic optimization problem, the classifier based on the support vector may be achieved.
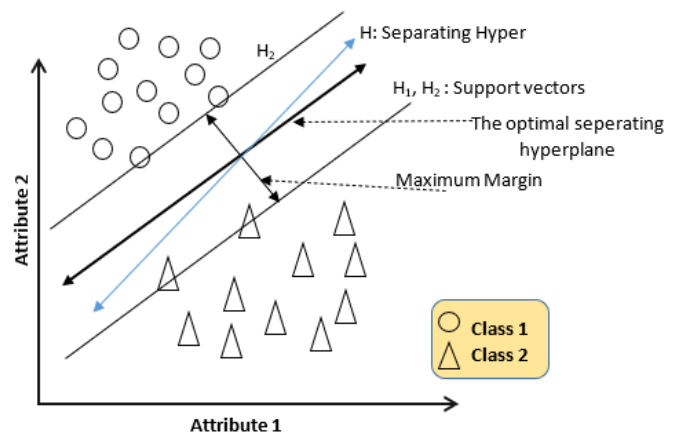

Figure 3. Classification of data points by SVM

After the training step, the SVM classifies a test sample using a decision function to determine on which side of the hyperplane the test sample lays. If those two classes are not linearly separable, it is possible to create an optimal separating

hyperplane that allows linear separation through a transformation that maps the data from input space to a high dimensional feature space by using a kernel function $K(x_i, x_j)$. In this work, the mapping is based on the radial basis function (RBF) kernel given by:

$$K(x_i, x_j) = exp\left(-\gamma \|x_i - x_j\|^2\right) \qquad (5)$$

where, γ is a parameter that controls the width of the Gaussian kernel. Here, $\gamma = 1$, the Matlab default value, is selected.

4. WAVELET FILTER DESIGN USING PARTICLE SWARM OPTIMIZATION (PSO)

4.1 Discrete wavelet transform

Over the last few decades, the wavelet transform has been widely and successfully applied to rotary machine fault diagnosis. Wavelets are functions constructed by translating and dilating a basic function called a mother wavelet $\psi(t)$:

$$\psi_{a,b}(t) = \psi\left(\frac{t - b}{a}\right) \qquad (6)$$

where, the parameter $a > 0$ represents the scale index (or dilation) and b indicates the time shifting (or translation).

The continuous wavelet transform (CWT) of a function $x(t) \in L^2(R)$ at scale a and position b can be described as a convolution with a scaled and translated copy of the mother wavelet $\psi(t)$ as:

$$CWT(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\psi^*\left(\frac{t - b}{a}\right) dt \qquad (7)$$

The asterisk * denotes the complex conjugation.

Since the scale and shift parameters are continuous, applying CWT on a signal generates information redundancy which results in an increased computation time. Thus, for applications requiring computational efficiency, the discrete wavelet transform (DWT) has been introduced:

$$DWT(j, k) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{+\infty} x(t)\psi\left(\frac{t - 2^j k}{2^j}\right) \qquad (8)$$

where, $a = 2^j$ and $b = k2^j$ represent a dyadic sampling of the dilation and translation parameters respectively. $(j, k) \in \mathbb{Z}^2$ are the decomposition level and the translation factor respectively. DWT is usually implemented using the iterative and efficient Mallat multiresolution algorithm through the use of a pair of a low-pass (LP) and high-pass (HP) digital filters, whose impulse response are respectively denoted by $h(n)$ and $g(n) = (-1)^n h(1 - n)$. Known as quadrature mirror filters (QMF), these two wavelets filters are often designed using the mother wavelet $\psi(t)$ and its corresponding scaling function $\phi(t)$ as [15, 18]:

$$\begin{cases} \phi(t) = \sqrt{2} \sum_n h(n)\phi(2t - n) \\ \psi(t) = \sqrt{2} \sum_n g(n)\phi(2t - n) \end{cases} \qquad (9)$$

Practically, the N_d-level DWT decomposition of a sampled vibration signal $x(t) = (x_1, x_2, x_3, \dots, x_N)$ is computed by passing the signal through N_d low pass (LP) h(n) and high pass (HP) g(n) filters, resulting in one approximation coefficient vector CA_{Nd} and N_d detail coefficient vectors CD_j respectively (j: decomposition level; $1 \le j \le N_d$). This multiresolution analysis, illustrated in Figure 4 for three-level decomposition, may be expressed by:
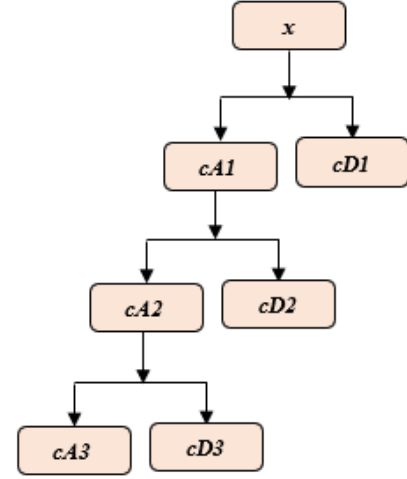

Figure 4. Three-level DWT decomposition tree

$$\begin{aligned} x(t) &= CA_{Nd}(t) + \sum_{j=1}^{N_d} CD_j(t) \\ &= \sum_n a_{N_d,n} \phi(t - n) + \sum_{j=1}^{N_d}\sum_n d_{j,n}\psi(2^j t - n) \end{aligned} \qquad (10)$$

Using the two wavelet filters h(n) and g(n), it can be shown that the signal may be recursively decomposed into low and high-frequency components [18, 28]:

$$\begin{cases} a_{m,n} = \sum_k h[k - 2n]a_{m-1,n} \\ d_{m,n} = \sum_k g[k - 2n]a_{m-1,n} \end{cases} \qquad (11)$$

where, $a_{m,n}$ and $d_{m,n}$ are respectively the approximation and detail coefficients of the multiresolution decomposition, obtained by convolving the previous level approximation coefficients $a_{m-1,n}$ with the LP and HP digital filter coefficients, respectively. For a N_d-level DWT decomposition, the total energy of the signal x(t), based on Parseval's theorem, is the energy $EA_{Nd} = \sum_n |a_{K,n}|^2$ of the approximation coefficient CA_{ND} plus the energies $ED_j = \sum_n |d_{j,n}|^2$ of the N_d detail coefficients CD_j:

$$E_{tot} = EA_{Nd} + \sum_{j=1}^{N_d} ED_j \qquad (12)$$

Note that, in this work, the energies EA_{Nd} and ED_j as well as the kurtosis values of the approximation and detail coefficients are employed as input features in the bearing fault classification process [17].

4.2 Wavelet filter bank and polyphase representation

The discrete wavelet transform can be successfully implemented using a pair of low-pass and high-pass finite impulse response (FIR) filters [18, 29]. The authors [19] proposed a recursive and interesting procedure to generate the coefficients of such filters. Based on polyphase matrix representation, this iterative algorithm allows the generation of orthonormal perfect reconstruction FIR filters of any given length. Let the transfer function, in the $\mathbb{Z}$ domain, of a low-pass FIR filter of order $2N_o$ given by:

$$H_0(z) = \sum_{i=0}^{2N_o-1} h_i \, z^{-i} \qquad (13)$$

Separating the even and odd index coefficients gives:

$$\begin{aligned} H_0(z) &= \sum_{i=0}^{N_o-1} h_{2i} z^{-2i} + z^{-1} \sum_{i=0}^{N_o-1} h_{2i+1} z^{-2i} \\ &= H_{00}(z^2) + z^{-1} H_{01}(z^2) \end{aligned} \qquad (14)$$

where, $H_{00}(z) = \sum_{i=0}^{N-1} h_{2i} z^{-i}$ and $H_{01}(z) = \sum_{i=0}^{N-1} h_{2i+1} z^{-i}$

Thus, the polyphase matrix of a two-channel filter bank is:

$$H_{poly}(z) = \begin{pmatrix} H_{00}(z) & H_{01}(z) \\ H_{10}(z) & H_{11}(z) \end{pmatrix} \qquad (15)$$

where, $H_{00}(z)$ and $H_{01}(z)$ are the polyphase components of the low-pass filter, whereas $H_{10}(z)$ and $H_{11}(z)$ are those of the high-pass filter. Vaidyanathan proposed the following polyphase matrix factorization [20]:

$$H_{poly}(z) = \begin{pmatrix} c_0 & s_0 \\ -s_0 & c_0 \end{pmatrix} \prod_{i=1}^{N_o-1} \begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix} \begin{pmatrix} c_i & s_i \\ -s_i & c_i \end{pmatrix} \qquad (16)$$

where, $c_i = cos(\theta_i)$ and $s_i = sin(\theta_i)$.

Sherlock et al. [19] developed a new reformulation of the factorization which leads to a recursive computation of the LP filter coefficients. Hence, the even numbered filter coefficients $\{h_{2i}\}$ are given by:

$$\begin{cases} h_0^{(k+1)} &= c_k h_0^k \\ h_{2i}^{(k+1)} &= c_k h_{2i}^{(k)} - s_k h_{2i-1}^{(k)} \quad i = 1,2,\dots,k-1 \\ h_{2k}^{(k+1)} &= -s_k h_{2k-1}^{(k)} \end{cases} \qquad (17)$$

with $h_0^{(1)} = c_0$ and $h_1^{(1)} = s_0$, while the odd filter coefficients $\{h_{2i+1}\}$ are defined by:

$$\begin{cases} h_1^{(k+1)} &= s_k h_0^{(k)} \\ h_{2i+1}^{(k+1)} &= s_k h_{2i}^{(k)} + c_k h_{2i-1}^{(k)} \quad i = 1,2,\dots,k-1 \\ h_{2k+1}^{(k+1)} &= c_k h_{2k-1}^{(k)} \end{cases} \qquad (18)$$

For the filters of length $2N_o$, Eq. (17) and Eq. (18) determine the low-pass coefficients $\{h_0, h_1, \dots, h_{2N_o-1}\}$ in terms of N_o randomly selected angular parameters

$\{\theta_0, \theta_1, \dots, \theta_{N_o-1}\}$ taking values in the interval $[0, \, 2\pi[$.

By flipping and sign-alternating, the HP filter coefficients are found by:

$$g_i = (-1)^{i+1} h_{2N_o-1-i} \qquad (19)$$

4.3 Particle Swarm Optimization

Particle Swarm Optimization (PSO) algorithm, introduced in 1995 by Kennedy and Eberhart [30], is a stochastic search technique inspired from the collective intelligent behavior of animal societies such as ant colonies, bird flocks, or fish schools. PSO aims at finding the optimal solution for a given multidimensional problem, where each candidate solution, called a particle, representing a point in the N- dimensional search space, where N is number of variables of the function to be optimized. PSO performs search starting from a random swarm of particles, called initial population, that emulate the success of neighboring members to achieve their own successes. Alternatively, using shared information amongst the swarm of individuals, each particle updates its velocity and position based on its current and past trajectories. Therefore, for a problem of N real variables to be optimized, and a population of size S, each particle $p_i (i = 1,2,\dots,S)$, at iteration t, has its own:

- Position $p_i(t) \in \mathcal{R}^N$.
- Velocity $v_i(t) \in \mathcal{R}^N$.
- Personnel best position $P_{bi}(t) \in \mathcal{R}^N$, found during past search.

Consider $P_g(t) \in \mathcal{R}^N$ being the best global position reached among the past trajectories explored by all the swarm members. During the search process, the particle movement may be described by [29]:

$$\begin{aligned} v_i(t+1) &= w \, v_i(t) + c_1.r_1[P_{bi}(t) - p_i(t)] \\ &\quad + c_2.r_2[P_g(t) - p_i(t)] \\ p_i(t+1) &= p_i(t) + v_i(t) \end{aligned} \qquad (20)$$

where, w is the inertia weight, used to provide a fine adjustment between local and global swarm searching capabilities, such that a greater value of w allows wider research area exploration. The constants c_1 and c_2 represent the cognition and the social learning factors regulating the relative velocities with respect to the best local and global positions, respectively. r_1 and r_2 are random variables taking values from a uniform distribution in the interval [0,1]. Eq. (20) is iterated until convergence is achieved. The PSO implementation is so simple that it is widely applied in many optimization problems. In this work, the particle swarm optimization algorithm is employed to search for the suitable angular parameters $\theta_i's$ related to the polyphase representation of wavelet filters in order to improve the wavelet-based bearing fault classification performance.

4.4 Wavelet filter optimization procedure

Wavelets, principally developed for general signal processing purpose, have been extensively and effectively applied in various vibration signal analysis problems, including faulty bearing diagnosis. Also, it has been shown that a better fault classification is achieved when using features extracted from wavelet transform. It is thought that the

classification performance could be boosted further if the wavelets are designed according to the classification accuracy. For this reason, based on the polyphase representation described in section 4.2 wavelet filters are designed using the PSO algorithm, where the positions of the swarm individuals encode the free parameters θ_i, and the bearing fault classification error rate is adopted as a cost function. Thus, the suggested optimization procedure, depicted in Figure 5 for a fixed filter order and a given decomposition level, is implemented as follows:

(1) Set the wavelet filter order $N_o = order\_min$.

(2) Set the decomposition level $N_d = decomp\_min$.

(3) Generate an initial population of particles, where each particle is an angular position vector $\{\theta_0, \theta_1, \dots, \theta_{N_0}\}$ randomly selected in the range $[0, 2\pi[$.

(4) For each particle evaluate the fitness function:
   a. Insert the $\theta_i's$ vector into the iterative algorithm to produce the $2N_o + 2$ coefficients $\{h_i\}$ of the low-pass filter;
   b. Determine the high-pass filter coefficients $\{g_i\}$ using the alternating flip algorithm given by Eq. (19).
   c. Using the obtained low-pass and high-pass filters apply the $N_d - level$ DWT to the vibration signal to extract the fault feature vector made up of the energies and the kurtosis:
   d. $[E_1, E_2, \dots, E_{N_d+1}, K_1, K_2, \dots, K_{N_d+1}]$ Insert the feature input vector into the classifier and evaluate the error rate percentage of the classification.

(5) Update the velocity and position of each particle using Eq. (20).

(6) Repeat steps 4-5 until convergence conditions are satisfied. Note that when there is no more improvement in the classification error rate percentage during 20 consecutive iterations, convergence is considered.

(7) Increment the decomposition level $N_d \leftarrow N_d + 1$, and repeat steps 3-6.

(8) Increment the filter order $N_o \leftarrow N_o + 2$, and repeat steps 2-7.

5. RESULTS AND DISCUSSIONS

To evaluate the performance of the considered wavelet filter optimization method and to assess the effectiveness of the extracted wavelet features with respect to the classification accuracy, several simulation tests employing the database described in section 2 have been carried out using the four previously presented classifiers, namely K-Nearest Neighbors (KNN), Fuzzy K-Nearest Neighbors (FKNN), Artificial Neural Network (ANN), and Support Vector Machine (SVM). Note that, the database, composed of 750 vibration signals, has been randomly partitioned into training data (70%) and testing data (30%). Generally, the particle swarm size is a

compromise between algorithm performance and consuming time [31]. Also, the maximum number of iterations is an important parameter that can affect the PSO algorithm performances, where a value between 50 and 200 is an appropriate suggestion that may lead to good results [32]. In this paper, the particle swarm size is taken to be 25 and the maximum number of generations is set to 100. To obtain good results, a linear decreasing inertia weight is adopted, and the cognitive and social parameters are set to $c_1 = 1.5$ and $c_2 = 2.0$ respectively [32]. To avoid the effect of initial population choice, 20 simulations are executed, each time, and the best results are saved. For a purpose of comparison, a thorough research, using the four classifiers, has shown that the Daubechies wavelet with the 3rd level decomposition performs relatively better compared to other wavelet families as shown in Table 2 and Figure 6.

Using the design steps described in section 4.4 with the 3-level decomposition and the four adopted classifiers, the obtained results, given in Table 3 and Figure 7, show that the designed wavelet filters give the best classification performance at filter orders greater or equal to 6. Hence, a wavelet filter order of 6 would represent an optimal choice.

It is noteworthy that the obtained FDD system is a scheme based on 3-level DWT decomposition, whose PSO-optimized filters of order 6 are reported in Table 4 for the four classifiers.

Also, the percentage values reported in Table 5 and represented in Figure 8, show clearly that the optimized wavelet filters enhance the classification accuracy with an average percentage gain of about 2%.

Table 2. Average classification accuracy (%) using the *"db"* wavelets in terms of the decomposition level for the four classifiers

Decomposition level N_d	KNN	FKNN	ANN	SVM
1	93.77	93.77	92.89	95.11
2	92.36	94.22	90.67	95.55
3	95.55	95.11	96	96.95
4	95.11	94.22	91.11	92.88
5	89.98	95.11	91.11	91.11
6	92.44	96.33	89.78	93.44
7	90.67	95.89	87.56	95.1
8	88.67	94.44	90.67	95.23

Table 3. Average classification accuracy (%) using the PSO-optimized wavelets versus the filter order for the four classifiers

	Designed wavelet Accuracy (%)			
Filter order	KNN	FKNN	ANN	SVM
4	97.33	98.22	96.89	97.33
6	97.47	98.87	97.78	98.40
8	97.47	98.87	97.78	98.40
10	97.47	98.87	97.78	98.40
12	97.47	98.87	97.78	98.40

Table 4. Designed wavelet filter coefficients of order 6

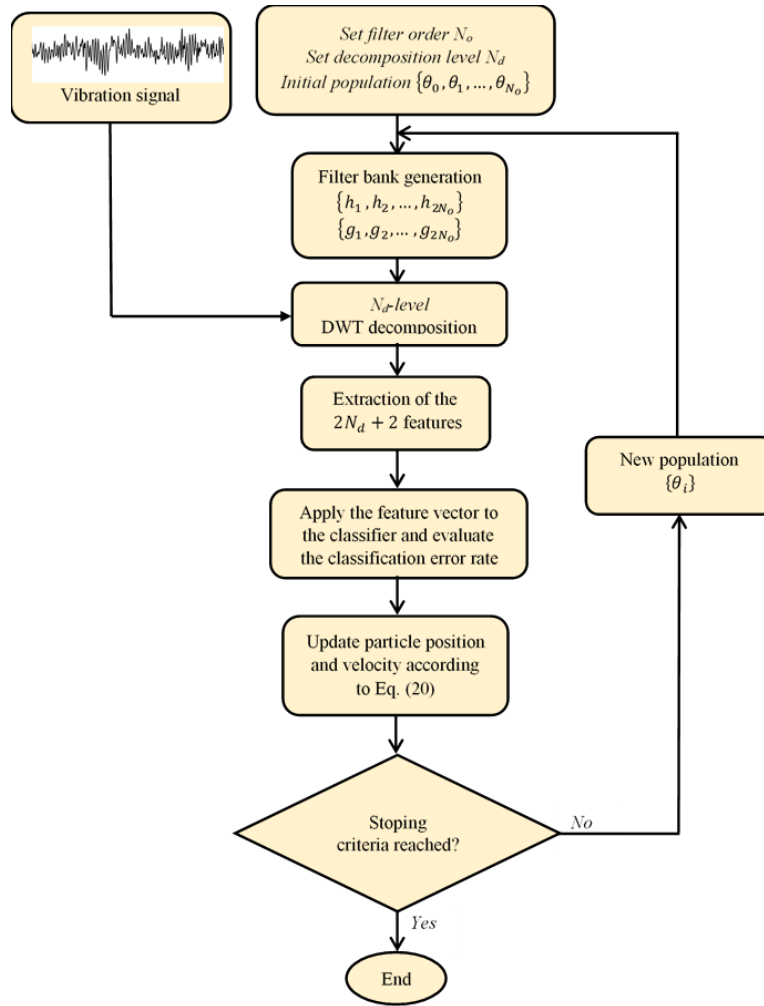|	Filter coefficients					
KNN	0.3828	0.1919	0.6180	-0.0707	-0.2937	0.5859
FKNN	0.6097	0.3238	-0.0462	0.6537	0.1436	-0.2704
ANN	0.8507	0.4723	-0.1191	0.1907	-0.0244	0.0440
SVM	0.4407	0.0038	0.2650	0.8456	0.0012	0.1424

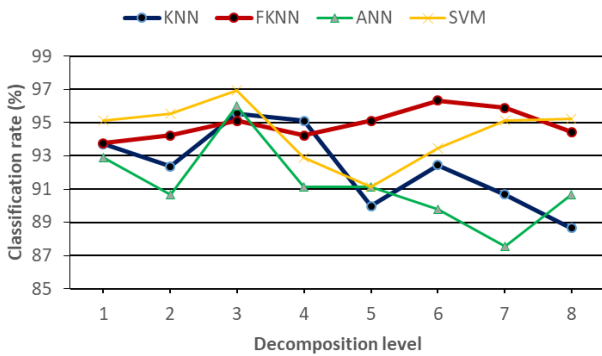**Figure 5.** Flowchart of PSO based design of the optimal wavelet filters

Figure 6. Classification rates using the *"db"* wavelets in terms of the decomposition level for the four classifiers
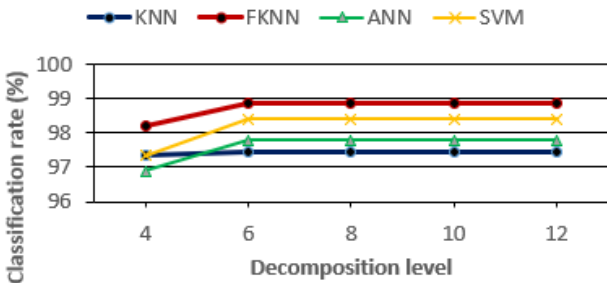

Figure 7. Classification rates using the ***designed*** wavelets in terms of the filter order for the four classifiers

Table 5. Bearing fault classification accuracy percentages

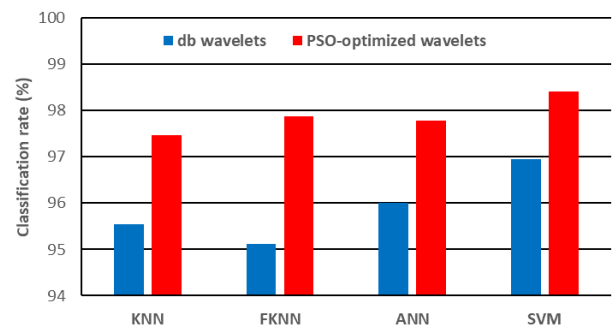|	KNN	FKNN	ANN	SVM
db **wavelets**	95.55	95.11	96	96.95
PSO-optimized wavelets	97.47	97.87	97.78	98.4

Figure 8. Comparison in terms of classification percentage between *db* and designed wavelets using the four classifiers

6. CONCLUSIONS

In this paper, a novel ball bearing fault diagnosis scheme based on the optimal design of discrete wavelet filters through the use of the polyphase representation of the wavelet filter bank is proposed. In order to assess the performance of the considered diagnosis technique, four popular classifiers with a

database consisting of 750 vibration signals representing five different states of the bearing, are employed. Some conclusions and contributions of this research are summarized as below.

(1) The new suggested technique is based on the design of wavelet filters, best adapted to bearing fault diagnosis, using the polyphase wavelet filter representation, and optimized using the PSO algorithm with respect to the bearing fault classification accuracy.

(2) By applying the PSO-designed wavelet filters to the vibration signal, two statistical features, the energy and kurtosis, are extracted and fed to four popular classifiers namely KNN, Fuzzy-KNN, ANN and SVM.

(3) Through a thorough investigation, it has been shown that the Daubechies *'db'* wavelets with a decomposition level $N_d = 3$, outperforms almost all the other standard discrete wavelets.

(4) Using the PSO optimized wavelet filters of order 6 for preprocessing the vibration signal, the developed approach is compared against the *'db'* based diagnosis system. The obtained simulation results show that the proposed design procedure results in optimal filters that surpasses the standard wavelet filters regardless the employed classifier, with an average classification percentage gain that amount to nearly 2%.

ACKNOWLEDGMENT

REFERENCES

[1] Tyagi, S., Panigrahi, S.K. (2017). A DWT and SVM based method for rolling element bearing fault diagnosis and its comparison with Artificial Neural Networks. Journal of Applied and Computational Mechanics, 3(1): 80-91. https://doi.org/10.22055/jacm.2017.21576.1108

[2] Kateris, D., Moshou, D., Pantazi, X.E., Gravalos, I., Sawalhi, N., Loutridis, S. (2014). A machine learning approach for the condition monitoring of rotating machinery. Journal of Mechanical Science and Technology, 28(1): 61-71. https://doi.org/10.1007/s12206-013-1102-y

[3] Vijay, G.S., Srinivasa Pai, P., Sriram, N.S. (2011). Discrete wavelet transform for fault diagnosis of rolling element bearing. National Conference on Advances in Mechanical Engineering, Manipal Institute of Technology, Manipal, pp. 3-5.

[4] Li, H., Liu, T., Wu, X., Chen, Q. (2019). Application of EEMD and improved frequency band entropy in bearing fault feature extraction. ISA Transactions, 88: 170-185. https://doi.org/10.1016/j.isatra.2018.12.002

[5] Chen, K., Huang, C., He, J. (2016). Fault detection, classification and location for transmission lines and distribution systems : A review on the methods. IET High Voltage, 1: 25-33. https://doi.org/10.1049/hve.2016.0005

[6] Zhang, X., Liu, Z., Wang, J., Wang, J. (2019). Time – frequency analysis for bearing fault diagnosis using multiple Q-factor Gabor wavelets. ISA Transactions, 87:

225-234. https://doi.org/10.1016/j.isatra.2018.11.033

[7] Heidari, M., Engineering, M., Branch, A. (2018). Using wavelet support vector machine for fault diagnosis of gearboxes. International Journal of Automotive Engineering, 8(1): 2604-2613. https://doi.org/10.22068/ijae.8.1.2603

[8] Kumar, A., Kumar, R. (2019). Role of signal processing, modeling and decision making in the diagnosis of rolling element bearing defect : A review. Journal of Nondestructive Evaluation, Springer US. 38(1): 1-29. https://doi.org/10.1007/s10921-018-0543-8

[9] Berrezzek, F., Khelil, K., Bouadjila, T. (2019). Efficient wind speed forecasting using discrete wavelet transform and artificial neural networks. Revue d'Intelligence Artificielle, 33(6): 447-452. https://doi.org/10.18280/ria.330607

[10] Khelil, J., Ramdani, M., Khelil, K., Boutasseta, N. (2019). Bearing faults diagnosis using discrete wavelets and artificial intelligence approaches. 1st International Conference on Sustainable Renewable Energy Systems and Applications (ICSRESA), Tebessa, Algeria.

[11] Chennippan, M., Bhaskaran, P.E., Adhulrasheed, I.S.K., Subramaniam, T., Govindasamy, R. (2020). Vibration signals based bearing defects identification through online monitoring using LABVIEW. Journal Européen Des Systèmes Automatisés, 53(2): 187-193. https://doi.org/10.18280/jesa.530204

[12] Shan, Y.X. (2018). Research of fault diagnosis of rolling bearing based on wavelet packet and FSVM research of fault diagnosis of rolling bearing based on wavelet packet and FSVM. Journal of Physics: Conference Series, 1087: 042061. https://doi.org/10.1088/1742-6596/1087/4/042061

[13] Kumar, H.S., Srinivasa, P.P., Sriram, N.S., Vijay, G.S. (2013). ANN based evaluation of performance of wavelet transform for condition monitoring of rolling element bearing. Procedia Engineering, 64: 805-814. https://doi.org/10.1016/j.proeng.2013.09.156

[14] Mohammed, O.A., Abed, N.Y., Ganu, S. (2006). Modeling and characterization of induction motor internal faults using finite-element and discrete wavelet transforms. IEEE Transactions on Magnetics, 42(10): 3434-6. https://doi.org/10.1109/TMAG.2006.879091

[15] Yan, R., Gao, R.X., Chen, X. (2014). Wavelets for fault diagnosis of rotary machines: A review with applications. Signal Processing, 96(PART A): 1-15. https://doi.org/10.1016/j.sigpro.2013.04.015

[16] Huo, Z., Zhang, Y., Francq, P., Shu, L., Huang, J. (2017). Incipient fault diagnosis of roller bearing using optimized wavelet transform based multi-speed vibration signatures. IEEE Access, 5: 19442-56. https://doi.org/10.1109/ACCESS.2017.2661967

[17] Patil, A.B., Gaikwad, J., Jayant, V.K. (2016). Bearing fault diagnosis using discrete wavelet transform and artificial neural network. 2nd International Conference on Applied and Theoretical Computing and Communication Technology (ICATccT), pp. 399-405. https://doi.org/10.1109/ICATCCT.2016.7912031

[18] Mallat, S.G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(7): 674-693. https://doi.org/10.1109/34.192463

[19] Sherlock, B.G., Monro, D.M. (1998). On the space of

orthonormal wavelets. IEEE Transactions on Signal Processing, 46(6): 1716-20. https://doi.org/10.1109/78.678504

[20] Vaidyanathan, P.P. (1993). Multirate Systems and Filter Banks. Prentice-Hall, Englewood Cliffs, NJ.

[21] Khelil, K., Berrezzek, F., Bouadjila, T. (2020). GA-based design of optimal discrete wavelet filters for efficient wind speed forecasting. Neural Computing and Applications, 1-14. https://doi.org/10.1007/s00521-020-05251-5

[22] Kotsiantis, S.B., Zaharakis, I.D., Pintelas, P.E. (2006). Machine learning: A review of classification and combining techniques. Artificial Intelligence Review, Springer. 26(3): 159-190. https://doi.org/10.1007/s10462-007-9052-3

[23] Keller, J.M., Gray, M.R., Givens, J.A. (1985). A fuzzy k-nearest neighbor algorithm. IEEE Transactions on Systems, Man, and Cybernetics, 15(4): 580-585. https://doi.org/10.1109/TSMC.1985.6313426

[24] Chen, H.L., Huang, C.C., Yu, X.G., Xu, X., Sun, X., Wang, G., Wang, S.J. (2013). An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach. Expert Systems with Applications, 40(1): 263-271. https://doi.org/10.1016/j.eswa.2012.07.014

[25] Liu, R., Yang, B., Zio, E., Chen, X. (2018). Artificial intelligence for fault diagnosis of rotating machinery: A review. Mechanical Systems and Signal Processing, Elsevier, 108: 33-47. https://doi.org/10.1016/j.ymssp.2018.02.016

[26] Widodo, A., Yang, B.S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. Mechanical Systems and Signal Processing, 21(6): 2560-2574. https://doi.org/10.1016/j.ymssp.2006.12.007

[27] Cristianini, N., Shawe-Taylor, J. (2000). An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press.

[28] Addison, P.S. (2017). The Illustrated Wavelet Transform Handbook. CRC Press.

[29] Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. Communications on Pure and Applied Mathematics, Wiley Online Library, 41(7): 909-996.

[30] Parsopoulos, K.E., Vrahatis, M.N. (2010). Particle swarm optimization and intelligence: Advances and applications. Information Science Reference, New York.

[31] Cheng, Y., Wang, Z., Zhang, W., Huang, G. (2019). Particle swarm optimization algorithm to solve the deconvolution problem for rolling element bearing fault diagnosis. ISA Transactions, 90: 244-267. https://doi.org/https://doi.org/10.1016/j.isatra.2019.01.012

[32] Tyagi, S., Panigrahi, S.K. (2017). An improved envelope detection method using particle swarm optimisation for rolling element bearing fault diagnosis. Journal of Computational Design and Engineering, 4(4): 305-317. https://doi.org/http://dx.doi.org/10.1016/j.jcde.2017.05.002

NOMENCLATURE

$a_{m,n}$	Approximation coefficients of the multiresolution decomposition
h_i	Low pass filter coefficients
CA_{Nd}	DWT approximation coefficient at decomposition level N_d
EA_{Nd}	Energy of the approximation coefficient
ED_j	Energy of the detail coefficient
N_d	DWT decomposition level
N_o	Filter order
$d_{m,n}$	Detail coefficients of the multiresolution decomposition
g_i	High pass filter coefficients
p_i	The PSO swarm particle
x_j	Sample vector
K	Number of nearest neighbors
C	fuzzy partition of C classes
CD_j	DWT detail coefficient at decomposition level j
θ_i	Angular parameters
μ_i	Membership degree of x_i
γ	Gaussian kernel control parameter