

Feature extraction and optimized support vector machine for severity fault diagnosis in ball bearing

Tawfik Thelaidjia^{abc*}, Abdelkrim Moussaoui^b and Salah Chenikher^c

^aDepartment of Electrical Engineering and Automatic, University of 08 Mai 1985, Guelma University, Algeria

^bLaboratory of Electrical Engineering of Guelma, Guelma University, Algeria

^cLaboratory of Electrical Engineering, LABGET, Tebessa University, Algeria

ARTICLE INFO

Article history:

Received 6 March, 2016

Accepted 30 June 2016

Available online

30 June 2016

Keywords:

Fault Diagnosis

Particle Swarm Optimization with

Passive Congregation

Principal Component Analysis

Statistical Parameters

Support Vector Machine

Wavelet Packet Transform

ABSTRACT

In this paper, a method for severity fault diagnosis of ball bearings is presented. The method is based on wavelet packet transform (WPT), statistical parameters, principal component analysis (PCA) and support vector machine (SVM). The key to bearing faults diagnosis is features extraction. Hence, the proposed technique consists of preprocessing the bearing fault vibration signal using statistical parameters and energy obtained through the application of Db8- WPT at the third level of decomposition. After feature extraction from vibration signal, PCA is employed for dimensionality reduction. Finally, particle swarm optimization with passive congregation-based support vector machine is used to classify seven kinds of bearing faults. The classification results indicate the effectiveness of the proposed method for severity faults diagnosis in ball bearings.

© 2016 Growing Science Ltd. All rights reserved.

1. Introduction

Rotating machinery is used in a wide variety of industrial applications, including aircraft engines, wind turbine generators and automobile transmission systems. Due to the growing of production capabilities in manufacturing process, machines must work continuously for extended hours. Consequently, undesired shutdowns due to machine breakdowns have become more expensive than ever before. Rolling bearings are critical components in rotating machinery, they represent the primary cause of machine defects (Bell et al., 1985; Zhou et al., 2007; Zarei et al., 2014). Therefore, bearing fault diagnosis is needed to avoid serious damage and economic losses. Signal processing is the first key of fault diagnosis systems, it is widely performed by combining several techniques including time, frequency and time-frequency domain (Lei et al., 2007; Dong & Luo, 2013; Pandya et al., 2014). By

* Corresponding author.

E-mail addresses: bilel.toufik@yahoo.fr (T. Thelaidjia)

combining a various kinds of features, we create a feature vector, which contains as much characteristic as possible about the bearing condition, thus a more comprehensive information is generated from the vibration signal. To acquire more fault characteristic information and improve the performance of the diagnosis system, both statistical parameters and energy obtained through the application of wavelet packet transform (WPT) are used for feature extraction. After extract the original features, it is difficult to select salient features, as different factors influence their effectiveness such as location of sensors, the noise due to the data acquisition system, etc (Dong & Luo, 2013; Malhi & Gao, 2004). However, the dimensionality of the obtained features needs to be reduced to decrease the complexity of the classification, so it is indispensable to select the useful features through redundancy reduction and removing insensitive features. Principal component analysis (PCA) is widely employed in feature extraction to obtain relevant information from multi-dimensional datasets (Dong & Luo, 2013; Saidi et al., 2015; Zhang et al., 2013; Zhou et al., 2014).

Classification is one of the most important tasks in faults diagnosis. different method have been investigated such as; neural network, ANFIS and SVMs. Support vector machine (SVM) is a new machine learning method; introduced by Vapnik and based on statistical learning theory, it has a welldefined formulation and is consistent with mathematical theory (Vapnik, 2013, 1998; Wang, 2005). Different from other classification methods, SVM does not require a large number of samples for solving classification problems (Chen et al., 2014; Zhi-qiang et al., 2005).

As the extra parameters of SVM viz: regularization and kernel parameters, play a crucial role in constructing an efficient classification model, it is an essential step to optimize these parameters. Inspired by the social behavior of fish schooling, bird flocking, and swarm theory, particle swarm optimization (PSO) has been widely used for SVM parameter optimization (Thelaidjia & Chenikher, 2013). Compared with other evolutionary computation techniques, PSO is easy for implementation and effective. Nevertheless it has the drawback that its performance may be decreased as the number of generations is increased (Angeline, 1998). In addition, while handling complex problems, PSO will easily converge towards local optimal solutions (Chen et al., 2014). To overcome these shortcomings, several variants have been developed to improve the PSO performance such as PSO with a constriction factor (CPSO) and Particle Swarm Optimizer with Passive Congregation (PSOPC) (Clerc & Kennedy, 2002; He et al., 2004). PSOPC has been successfully applied in many areas, it gives better results than those generated by other variants of PSO (He et al., 2004).

The paper is organized as follows. Section 2 describes the feature extraction method based on statistical parameters and WPT followed by PCA. Sections 3 and 4 explain the SVM and PSOPC, respectively. Section 5 presents the experimental results.

2. Feature extraction

The vibration signals are firstly processed by statistical parameters and a WPT at the third level using DB8 as mother wavelet to construct a combined feature set.

2.1. Features in the time domain

Statistical time domain parameters, such as: variance, kurtosis, crest indicator, etc..., are effective and practical in bearing fault diagnosis due to their sensitivity to the change of the signal form (Stepanic et al., 2009). Thus to describe the vibration signal in the time domain, the following statistical parameters are used:

$$Min = \min(x), \quad (1)$$

$$PPV = [\max(x) - \min(x)], \quad (2)$$

$$RMS = \sqrt{\frac{\sum_{k=1}^K x^2(k)}{K}}, \quad (3)$$

$$SD = \sqrt{\frac{\sum_{k=1}^K (x(k) - \bar{x})^2}{K}}, \tag{4}$$

$$CRF = \frac{PPV}{RMS}, \tag{5}$$

$$IMF = \frac{PPV}{\frac{1}{K} \sum_{k=1}^K x(k)}, \tag{6}$$

$$Kur = \frac{\sum_{n=1}^N (x(k) - \bar{x})^4}{N\sigma^2}, \tag{7}$$

where $x(k)$ is the vibration signal for $k=1,2,\dots,K$, \bar{x} is the mean of the vibration signal and σ is the variance of the vibration signal.

2.2 Wavelet packet algorithm

In WPT details are also splitted into two sub-signals including a low frequency and a high frequency band. It is more effective for wide frequency band analysis compared with wavelet transform (WT)(Zhang et al., 2013; Chebil et al., 2009). The energy in each frequency band differs from a vibration signal to another one according to the bearing faults (Pandya et al., 2014; Chebil et al., 2009). The application of WPT requires the selection of two descriptive parameters viz: the mother wavelet and the decomposition level (Pandya et al., 2014; Djebala et al., 2015).

In this paper, five wavelets are considered for feature extraction (Daubechies wavelet (db5 , db8), Reverse bi orthogonal wavelet (rbio5.5), symlets wavelet (sym2) and coiflet wavelet (coif5)). For selcting the appropriate wavelet mother both Maximum relative wavelet energy (MRWE) and Maximum energy-to-Shannon entropy ratio criteria (MESER) are used (Pandya et al., 2014; Kankar et al., 2011a). The calculating results are shown in Fig. 1 and Fig. 2. From the results presented in Figs. 1 and 2, it can be seen that the db8 wavelet gives good results. Thus it is considered as the most appropriate wavelet for feature extraction using WPT.

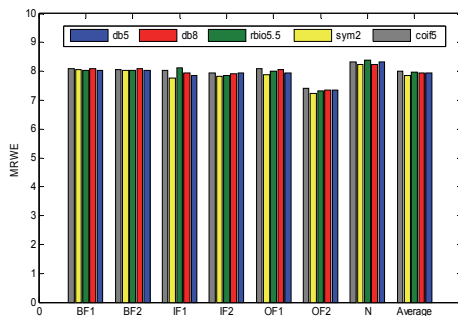


Fig. 1. Relative wavelet energy vs bearing condition using db5, db8, rbio5.5, sym2, coif5

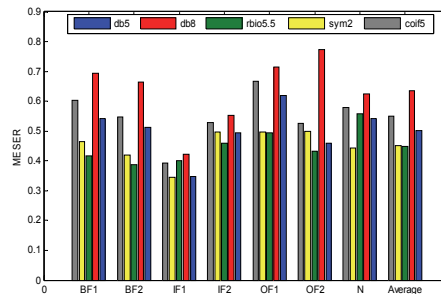


Fig. 2. Max. energy/Shannon entropy vs bearing condition using db5, db8, rbio5.5, sym2, coif5

In order to choose the decomposition level ,Vibration signals are decompose at the third level based on the followin equation (Djebala et al., 2015):

$$n \leq \log_2\left(\frac{F_{\max}(s)}{3F_c}\right), \tag{8}$$

where, n is the decompostion level, F_{\max} is the maximal frequency of the vibration signal and F_c is the defect frequency.

The step of feature extraction based on three layer wavelet packet is given as follows :

Firstly, the measured signal $x(t)$ was decomposed using db8 mother wavelet at the 3rd level. Thus, eight frequency components are obtained. Secondly, the decomposition coefficient (d_{ni}) of each frequency band are reconstructed and the corresponding signal is generated. D_{30} is the obtained signal from d_{30} , similarly D_{3i} , represent the reconstructed signal of d_{3i} .

The total signal is defined as :

$$D = D_{30} + D_{31} + D_{32} + D_{33} + D_{34} + D_{35} + D_{36} + D_{37} \quad (9)$$

Finally, The wavelet packet energy of each frequency band is expressed as :

$$E_{3j} = \sum_{i=1}^j |d_{ni}|^2 \quad (10)$$

where $n=3$;

Normalized Let,

$$T = \sum_{l=0}^7 E_{3l} \quad (11)$$

$$P_{3l} = \frac{E_{3l}}{T} \quad (12)$$

The relative energy Eigenvector E is defined as:

$$E = [P_{30} \ P_{31} \ P_{32} \ P_{33} \ P_{34} \ P_{35} \ P_{36} \ P_{37}] \quad (13)$$

2.3. Principal components analysis (PCA)

PCA constructs a set of uncorrelated multivariate samples called principal components (PC's) from an original correlated variable through an orthogonal transformation.

Let

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}$$

where n and p ($p=15$) represent the number of samples and the number of features, respectively. The steps of PCA to obtain a new uncorrelated set are as follows:

1. Calculating the correlation coefficient matrix

The correlation coefficient matrix is calculated using the following equation:

$$R = Cor(i, j) = \frac{(n-1) \cdot Cov(i, j)}{\sqrt{\sum_{k=1}^n (x_i(k) - \mu_i)^2 \sum_{k=1}^n (x_j(k) - \mu_j)^2}} = r(i, j)_{p \times n} \quad (14)$$

The dimension of the correlation matrix R is $(m \times n)$.

$Cov(i, j)$ is the covariance matrix $(p \times n)$ is and it is given by:

$$Cov(i, j) = \frac{1}{(n-1)} (x_i - \mu_i)(x_j - \mu_j); i, j = 1, 2, \dots, p \quad (15)$$

Here μ_i and μ_j are the averages of the i th and j th rows of X , respectively.

2. Calculating the eigenvectors V_i and eigenvalues λ_i from the correlation matrix R . The m eigenvalues are then ranked in descending order ($\lambda_1 > \lambda_2 > \dots > \lambda_m$).

λ_i and V_i satisfy the following condition:

$$A V_i = \lambda_i \cdot V_i ; i = 1, 2, \dots, p \quad (16)$$

where A is a $p \times n$ covariance matrix and V_i is given as below:

$$V_i = [V_{1i}, V_{2i}, \dots, V_{pi}]$$

3. Generates the new samples

the compressed eigenvectors are computed as follows:

$$X_{new} = V^T \cdot X, \quad (17)$$

where X and X_{new} represent the the original samples and the new uncorrelated samples. V is is called the weight matrix.

Finally, The first m ($m < p$) principal components in X_{new} according to the variance threshold (λ_i). are selected to form a new compressed feature vector.

3. Support vector machine (SVM)

SVM is a successful supervised algorithm which transform the original data with a low dimension into a high-dimensional feature space. As shown in Fig.3, SVM aims to find an optimal hyperplane by maximizing the margin between two datasets.

Given a dataset with n examples (x_i, y_i) , where each x_i is an observation and y_i its associated decision belong to $\{-1, +1\}$. By using a nonlinear mapping ϕ_x , the original data is mapped into a new feature space in which the data are sparse and possibly more separable and the maximum margin separating hyperplane $w\phi(x) + b_0$ will be built. Here b_0 is is an offset term.

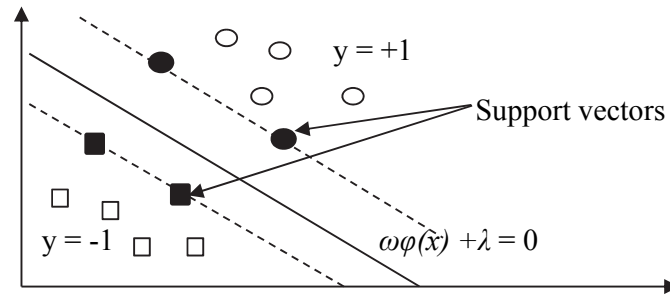


Fig. 3. Classification using SVM

In order to find the optimal hyperplane, we need to minimize the following functional (Vapnik, 1998; Wang, 2005).

$$\psi(w) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \quad (18)$$

subject to

$$y_k(\langle w; x \rangle - b_0) - 1 \geq 1 - \xi_i \quad (19)$$

where C is a regularization parameter which controls the tradeoff between the margin maximisation and error minimisation.

After solving the Lagrange equation of Eq. (18), the decision function is given as:

$$D(x) = \text{sgn}[\sum_{k \in N_{sv}} \alpha^* y_k \kappa(x_k^T x) - \lambda_0^*] \quad (20)$$

Here, α_i is the Lagrange coefficient and $\kappa(x_k, x)$ is called kernel function. In this paper, the popular radial basis function (RBF) is adopted and its mathematical formula is given by:

$$\kappa(x_1, x_2) = \exp(-\|x_1 - x_2\|^2 / 2\delta^2) \quad (21)$$

where parameter δ^2 is the variance of the Gaussian function.

Support vector machines can only deal with binary classifiers, several approaches have been designed to expand SVM for multi-classifying field (Abe, 2005). In this paper, we use the one-against-the-rest algorithm to realize a multi-faults classifier.

4. Particle Swarm Optimizer with passive congregation

PSO is a widely used metaheuristic algorithm which performs searches using a population of individuals. Each member of the population (particle), which is characterized by its current position (Z_{id}^{k-1}), its velocity (V_{id}^{k-1}) and its best position (P_{id}), is updated from an iteration to another one using its best previous position (P_{id}), and the best position among all the particles (P_{Gd}).

The update of each particle position is expressed as follow:

$$Z_{id}^k = Z_{id}^{k-1} + V_{id}^k \quad (22)$$

where:

$$V_{id}^k = w \cdot V_{id}^{k-1} + c_1 \cdot r_1 \cdot (P_{id} - Z_{id}^{k-1}) + c_2 \cdot r_2 \cdot (P_{Gd} - Z_{id}^{k-1}) \quad (23)$$

The PSOPC is an enhanced version of the PSO, it aims to increase the swarm diversity and to permit an escape from local minima using an additional part in the velocity update based on randomly selected particle from the swarm.

The update of each particle velocity is expressed as follow (He et al., 2004) :

$$V_{id}^k = w \cdot V_{id}^{k-1} + c_1 \cdot r_1 \cdot (P_{id} - Z_{id}^{k-1}) + c_2 \cdot r_2 \cdot (P_{Gd} - Z_{id}^{k-1}) + c_3 \cdot r_3 \cdot (P_R - Z_{id}^{k-1}) \quad (24)$$

in which, P_{Gd} : is the best position among all the particles, P_{id} : is the best position of the i^{th} particle, P_R : is a randomly selected particle from the swarm, Z_{id}^k : the position of the i^{th} particle at the k^{th} iteration, V_{id}^k : the velocity of the i^{th} particle at the k^{th} iteration, c_1, c_2, c_3 : are constants known as the cognitive, social and congregation acceleration coefficients, respectively. w : is the inertia weight and r_1, r_2, r_3 : are random numbers between 0 and 1.

The detailed procedure of PSOPC based- SVM is as follows:

Step 1. Initialize the population and set Parameters: Set the following PSOPC parameters: $(c_1, c_2, c_3) = (0.5, 0.5, 0.6)$, position ("C" and "σ" of SVM where $C \in [0; 1000]$, and $\delta \in [0; 1000]$) and velocity of each particle, the swarm size (set to 30) and the number of iterations (set to 60).

Step 2. Evaluate fitness: Evaluate the fitness function of each particle in the swarm and calculate the classification accuracy.

Step 3. Update P_i and P_G .

Step 4. Particle manipulations: Update the position of each particle using formula (22) and (24).

Step 5. Test of the termination condition: If the stopping condition is satisfied, stop the algorithm and save optimal "C" and "σ" for SVM, otherwise, go to **step 2**.

The inertia weight is given as follow:

$$w(k) = w_{max} - \frac{(w_{max} - w_{min})}{k_{max}} \cdot k \quad (25)$$

where: w_{max} and w_{min} represent the initial and the final weight, respectively. k_{max} represents the maximum number of generation and 'k' is the actual iteration number.

5. Analysis of experimental results

The vibration signals obtained from the Western Reserve University Bearing Data Center Website (Case Western Reserve University, 2014) are used in this study. The database is composed from seven (07) different classes, including a normal bearing (N) and six faults of ball bearing (bearings with ball fault of 0.1778mm (BF1), bearings with ball fault of 0.5334mm (BF2), bearings with inner race fault of 0.1778mm (IF1), bearings with inner race fault of 0.5334mm (IF2), bearings with outer race fault of 0.1778mm (OF1) and bearings with outer race fault of 0.5334mm (OF2)).

Forty samples are collected for each class, twenty six measurements are performed for training and fourteen for test purposes. The proposed diagnostic procedure is depicted in Fig. 4. First, the bearing vibration signal is preprocessed using statistical parameters and energy obtained through the application of three level WPT. As a result a combined feature set is constructed. PCA is then introduced for dimensionality reduction. Finally the obtained vector is used as input of a PSOCP-based SVM to achieve the classification step.

Fig. 5 and Fig. 6 show the distribution of the input training and test data. It can be seen that the input features are well clustered and ready for classification. Table 1 gives the classification results using a compressed feature set composed from the first two, three and four principal components. From Table 1 it can be seen that the first three principal components are sufficient to obtain good results. Part of Eigenvectors for different bearing conditions are given in Table 2.

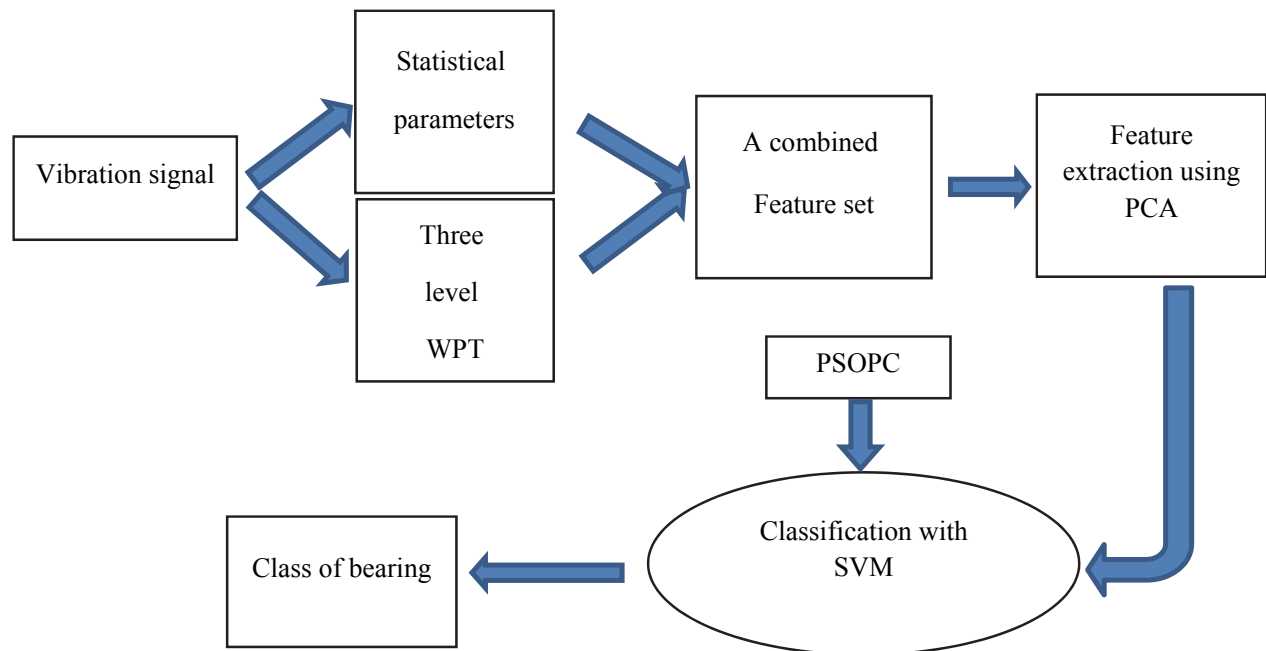


Fig. 4. The proposed methodology for bearing fault diagnosis

Table 3 show the classification performances for this severity fault classification problem using :

- The proposed method (PSOCP-SVM based on PCA feature extraction).
- PSOCP-SVM based on three feature sets selected randomly.
- PSOCP-SVM based on the combined feature set (15 features).

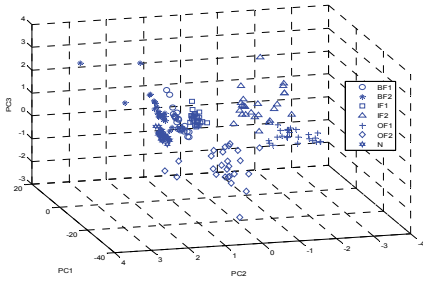


Fig. 5. Three PCs obtained from PCA (Training data)

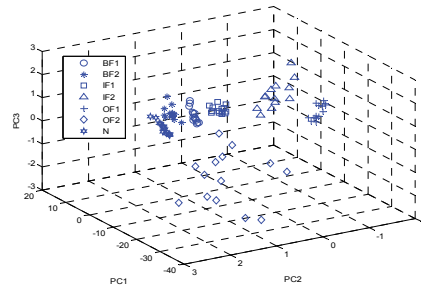


Fig. 6 Three PCs features obtained from PCA (Test data)

Table 1. SVM Classification Results of Bearing Fault for Different PCS

| Size of input after applying PCA | Optimal δ | Optimal C | Rate of validation (%) | Rate of test (%) |
|----------------------------------|------------------|-----------|------------------------|------------------|
| 2 | 0.9137 | 171.3802 | 98.35 | 98.98 |
| 3 | 1.0508 | 787.9622 | 100 | 100 |
| 4 | 0.9177 | 378.1020 | 100 | 100 |

Table 2. Part of Eigenvectors and Recognition Results

| Input | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|--------|--------|--------|--------|---------|---------|---------|---------|
| PC1 | 9.1141 | 5.9640 | 3.7421 | -1.7840 | -0.8955 | -23.722 | 10.1288 |
| PC2 | 0.4441 | 1.2340 | 0.0895 | -1.9907 | -2.8952 | 0.3143 | 0.9108 |
| PC3 | 0.0010 | 1.0766 | 0.5802 | 1.4671 | -0.8193 | -0.3941 | -0.6745 |
| State | N | BF1 | BF2 | IF1 | IF2 | OF1 | OF2 |
| Result | N | BF1 | BF2 | IF1 | IF2 | OF1 | OF2 |

Table 3. Classification results

| Feature Extraction | Traning rate(%) | Test rate(%) |
|--------------------------|-----------------|--------------|
| The proposed method | 100 | 100 |
| SD P_{32} IMF | 99.45 | 98.98 |
| RMS CRF Kur | 99.45 | 89.80 |
| Kur P_{31} Min | 100 | 95.92 |
| The combined feature set | 100 | 100 |

From experiments we can find that:

- The best accuracy for both validation (100%) and test set (100%) is obtained by using the proposed approach of SVM-PSOPC based on PCA feature extraction, where all the samples are correctly classified.
- When we use the randomly selected feature sets, the test classification accuracy decreases until 89.80%, it is difficult to select the feature set which gives the best results.
- When we use the global feature set (15 features), good classification rate is obtained, this is due to the capability of SVM to solve the complex classification problems.

To evaluate the efficiency of our method, a comparison between the proposed methodology and some previous studies is depicted in Table 4. The comparison has been done based on dimension of

the input feature, classifier used, nombres of the considered classes, and the classification accuracy. As a result presented data in this Table demonstrates that the proposed approach produces a good classification results.

6. Conclusion

In this paper, a new approach for severity faults diagnosis of rolling bearings is presented. In this method, we adopted statistical parameters and energy obtained through the application of WPT for feature extraction. Principal component analysis is then employed for dimensionality reduction. The resulting features have shown to be able to produce good input for classification. In addition PSOPC is successfully applied to optimize the SVM parameters. It can effectively extract the optimal parameters for building a good classification model. Results revealed that the proposed methodology based on wavelet packet transform, statistical parameters, principal component analysis and PSOPC-SVM is very effective and it can successfully diagnose the bearing condition.

Table 4. Comparison of this work with some previous studies

| Refrence | Nombre of features | Classifier used | Nombres of classes | Classification effeciency |
|----------------------|--------------------|--------------------|--------------------|---|
| Kankar et al., 2011b | 06 | ANN and SVM | 05 | Maximum classification rate is reported as 73.9726 % with SVM |
| Prieto et al., 2013 | 02 | MLP(4 classifiers) | 06 | Test rate is 95% |
| Sharma et al., 2016 | 07 | SVM and ANN | 07 | Maximum classification rate is reported as 100 % |
| Proposed methodology | 03 | SVM | 07 | Test rate is 100% |

Acknowledgement

The authors sincerely thank the Case Western Reserve University, U.S.A. for providing access to the bearing data sheet.

References

- Abe, S. (2005). *Support vector machines for pattern classification* (Vol. 53). London: Springer.
- Angeline, P. J. (1998, March). Evolutionary optimization versus particle swarm optimization: Philosophy and performance differences. In *International Conference on Evolutionary Programming* (pp. 601-610). Springer Berlin Heidelberg.
- Bell, R. N., McWilliams, D. W., O'donnell, P., Singh, C., & Wells, S. J. (1985). Report of large motor reliability survey of industrial and commercial installations. I. *IEEE Transactions on Industry applications*, 21(4), 853-864.
- Case Western Reserve University. (Last accessed in 2014), Bearing data center.
- Clerc, M., & Kennedy, J. (2002). The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *IEEE transactions on Evolutionary Computation*, 6(1), 58-73.
- Chen, F., Tang, B., Song, T., & Li, L. (2014). Multi-fault diagnosis study on roller bearing based on multi-kernel support vector machine with chaotic particle swarm optimization. *Measurement*, 47, 576-590.
- Chebil, J., Noel, G., Mesbah, M., & Deriche, M. (2009). Wavelet decomposition for the detection and diagnosis of faults in rolling element bearings. *Jordan Journal of Mechanical and Industrial Engineering*, 3(4), 260-267.

- Djebala, A., Babouri, M. K., & Ouelaa, N. (2015). Rolling bearing fault detection using a hybrid method based on empirical mode decomposition and optimized wavelet multi-resolution analysis. *The International Journal of Advanced Manufacturing Technology*, 79(9-12), 2093-2105.
- Dong, S., & Luo, T. (2013). Bearing degradation process prediction based on the PCA and optimized LS-SVM model. *Measurement*, 46(9), 3143-3152.
- He, S., Wu, Q. H., Wen, J. Y., Saunders, J. R., & Paton, R. C. (2004). A particle swarm optimizer with passive congregation. *Biosystems*, 78(1), 135-147.
- Kankar, P. K., Sharma, S. C., & Harsha, S. P. (2011). Fault diagnosis of ball bearings using continuous wavelet transform. *Applied Soft Computing*, 11(2), 2300-2312.
- Kankar, P. K., Sharma, S. C., & Harsha, S. P. (2011). Fault diagnosis of ball bearings using machine learning methods. *Expert Systems with Applications*, 38(3), 1876-1886.
- Lei, Y., He, Z., Zi, Y., & Hu, Q. (2007). Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs. *Mechanical Systems and Signal Processing*, 21(5), 2280-2294.
- Malhi, A., & Gao, R. X. (2004). PCA-based feature selection scheme for machine defect classification. *IEEE Transactions on Instrumentation and Measurement*, 53(6), 1517-1525.
- Pandya, D. H., Upadhyay, S. H., & Harsha, S. P. (2014). Fault diagnosis of rolling element bearing by using multinomial logistic regression and wavelet packet transform. *Soft Computing*, 18(2), 255-266.
- Prieto, M. D., Cirrincione, G., Espinosa, A. G., Ortega, J. A., & Henao, H. (2013). Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks. *IEEE Transactions on Industrial Electronics*, 60(8), 3398-3407.
- Saidi, L., Ali, J. B., & Fnaiech, F. (2015). Application of higher order spectral features and support vector machines for bearing faults classification. *ISA transactions*, 54, 193-206.
- Sharma, A., Amarnath, M., & Kankar, P. K. (2016). Feature extraction and fault severity classification in ball bearings. *Journal of Vibration and Control*, 22(1), 176-192.
- Stepanic, P., Latinovic, I. V., & Djurovic, Z. (2009). A new approach to detection of defects in rolling element bearings based on statistical pattern recognition. *The International Journal of Advanced Manufacturing Technology*, 45(1-2), 91-100.
- Thelaidjia, T., & Chenikher, S. (2013, December). A New approach of preprocessing with SVM optimization based on PSO for bearing fault diagnosis. In *Hybrid Intelligent Systems (HIS), 2013 13th International Conference on* (pp. 319-324). IEEE.
- Vapnik, V. (2013). *The nature of statistical learning theory*. Springer Science & Business Media.
- Vapnik, V.N. (1998). *Statistical Learning Theory*, Springer, New York.
- Wang, L. (Ed.). (2005). *Support vector machines: theory and applications* (Vol. 177). Springer Science & Business Media.
- Zarei, J., Tajeddini, M. A., & Karimi, H. R. (2014). Vibration analysis for bearing fault detection and classification using an intelligent filter. *Mechatronics*, 24(2), 151-157.
- Zhang, Z., Wang, Y., & Wang, K. (2013). Intelligent fault diagnosis and prognosis approach for rotating machinery integrating wavelet transform, principal component analysis, and artificial neural networks. *The International Journal of Advanced Manufacturing Technology*, 68(1-4), 763-773.
- Zhi-qiang, J., Hang-guang, F., & Ling-jun, L. I. (2005). Support Vector Machine for mechanical faults classification. *Journal of Zhejiang University Science A*, 6(5), 433-439.
- Zhou, W., Habetler, T. G., & Harley, R. G. (2007, September). Bearing condition monitoring methods for electric machines: A general review. In *Diagnostics for Electric Machines, Power Electronics and Drives, 2007. SDEMPED 2007. IEEE International Symposium on* (pp. 3-6). IEEE.
- Zhou, Z., Liu, D., & Shi, X. (2014). Fault Diagnosis Based on Principal Component Analysis and Support Vector Machine for Rolling Element Bearings. In *Practical Applications of Intelligent Systems* (pp. 795-803). Springer Berlin Heidelberg.