Engineering Science and Technology, an International Journal 17 (2014) 152–157

HOSTED BY
ELSEVIER

Contents lists available at ScienceDirect

Engineering Science and Technology, an International Journal

journal homepage: http://www.elsevier.com/locate/jestch

Full length article

Fault diagnosis of monoblock centrifugal pump using SVM

V. Muralidharan [a,*], V. Sugumaran [b], V. Indira [c]

[a] Department of Mechanical Engineering, B.S. Abdur Rahman University, Tamil Nadu, India
[b] School of Mechanical and Building Sciences, VIT University, Chennai Campus, Tamil Nadu, India
[c] Department of Mathematics, PKAC, Kalitheerthalkuppam, Puducherry, India

ARTICLE INFO

Article history:
Received 1 February 2014
Received in revised form
17 April 2014
Accepted 24 April 2014
Available online 29 May 2014

Keywords:
Mono-block centrifugal pump
SVM algorithm
Fault diagnosis
Continuous wavelet transforms (CWT)

ABSTRACT

Monoblock centrifugal pumps are employed in variety of critical engineering applications. Continuous monitoring of such machine component becomes essential in order to reduce the unnecessary break downs. At the outset, vibration based approaches are widely used to carry out the condition monitoring tasks. Particularly fuzzy logic, support vector machine (SVM) and artificial neural networks were employed for continuous monitoring and fault diagnosis. In the present study, the application of SVM algorithm in the field of fault diagnosis and condition monitoring is discussed. The continuous wavelet transforms were calculated for different families and at different levels. The computed transformation coefficients form the feature set for the classification of good and faulty conditions of the components of centrifugal pump. The classification accuracies of different continuous wavelet families at different levels were calculated and compared to find the best wavelet for the fault diagnosis of the monoblock centrifugal pump.

1. Introduction

As centrifugal pumps play a vital role in many critical applications, the continuous availability of such mechanical components become absolute essential. The pumps are the key elements in waste water treatment plants, food industry, agriculture, oil & gas industry, paper & pulp industry *etc*. In a monoblock centrifugal pump, the performance of bearing and impeller has direct effect on the desired pump characteristics. Faulty bearing, defect on the impeller and cavitation are main sources of many serious problems such as noise, high vibration *etc*. Cavitation can cause more undesirable effects, such as deterioration of the hydraulic performance (drop in head capacity and efficiency), damage of the pump by pitting, erosion and structural vibration. Vibration signals are widely used in condition monitoring of centrifugal pumps. Fault detection is achieved by comparing the signals and the similarities of monoblock centrifugal pump running under normal and faulty conditions. The faults considered in this study are bearing fault (BF), impeller fault (IF), bearing and impeller fault (BFIF) together and cavitation (CAV). In conventional condition monitoring, the vibration analysis is carried out with Fast Fourier Transform (FFT). With the help of seismic or piezoelectric transducers, the level of vibration can be measured. For complex systems involving many components, it is difficult to compute characteristic fault frequencies. Even if characteristic frequencies are available the vibration signals are highly non-stationary in nature and FFT based methods may not be suited for such processes. In the machine learning approach the data acquisition system is used to capture the vibration signals. From the vibration signal relevant features can be extracted and classified using a classifier. The step by step procedure for classification of faults is presented in Fig. 1.

H.Q. Wang et al., (2007) presented a fault diagnosis method for a centrifugal pump with frequency domain symptom parameter by using wavelet transform (feature extraction), rough sets (rule generation) and fuzzy neural network (classification) to detect faults and distinguish fault types at early stages [1].

V. Muralidharan and V. Sugumaran (2012) have reported the comparative performance of Naive Bayes and Bayes net algorithm for monoblock centrifugal pump. This paper mainly deals the flexibility of Bayes algorithm as a classifier [2]. V. Muralidharan, V. Sugumaran and N.R. Sakthivel (2011) proposed the application of SVM classifier for the decomposed wavelet features and also the classification performances. This study deals the decomposition of signals using discrete wavelet transforms [3,10]. Kemal Polat and Salih Gunes (2009) proposed a novel hybrid classification system based on J48 algorithm and one-against-all approach to classify the multi-class problems including dermatology, image segmentation

* Corresponding author.
E-mail address: v.muralidharan2@gmail.com (V. Muralidharan).
Peer review under responsibility of Karabuk University.

http://dx.doi.org/10.1016/j.jestch.2014.04.005

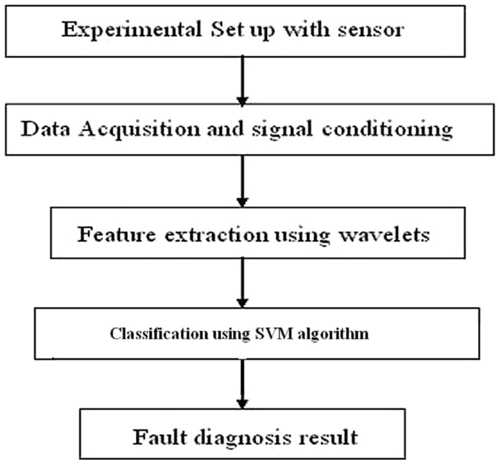

Fig. 1. Flow chart of monoblock centrifugal pump fault diagnosis system.

and lymphography. Initially, J48 algorithm has been used and achieved 84.48%, 88.79%, and 80.11% classification accuracies for dermatology, image segmentation, and lymphography datasets respectively. The proposed method based on J48 algorithm and one-against-all approach obtained 96.71%, 95.18%, and 87.95% for above datasets, respectively [4]. Muralidharan and Sugumaran (2013) reported the systematic methodology for wavelet selection using J48 algorithm. However, SVM based method has very strong mathematical background which may give more reliable results when compared to other algorithms [5,11].

Fansen Kong and Ruheng Chen (2004) proposed a new combined diagnostic system for triplex pump based on wavelet transform, fuzzy logic and neural network. The developed diagnostic system consists of four parts. The first part was wavelet transform for multiresolution analysis. The second part was for asymptotic spectrum estimation of the characteristic variable. The third part was employed for characteristic variable fuzzified in simulating fuzzy inference using incomplete information. The fourth part was the neural network trained with fuzzified characteristic variable for triplex pump failure diagnosis [6,12]. Jiangping Wang and Hangtao Hu (2006) used fuzzy logic principle as classifier with the features extracted from the vibration signals of the pump [7,9]. Javier Sanz et al., (2007) presented a technique for monitoring the condition of rotating machinery from vibration analysis that combines the capability of wavelet transform to treat the transient signals with the ability of auto associative neural networks to extract features of datasets in an unsupervised mode. Trained and configured networks with wavelet transform coefficients of non faulty signals were used as a method to detect the novelties or anomalies of faulty signals [8,13,14]. Also, V. Muralidharan and V. Sugumaran (2012) have illustrated the feature extraction using wavelets and classification using J48 algorithm for a fault diagnosis problem. In all of the above papers, researchers have reported with high classification accuracy. However, using Bayes algorithm and J48 algorithm requires a lot of domain expertise and computational time whereas the extraction of wavelet features and SVM classifiers are mathematically proved and validated with lots of bench marked complex datasets. Hence, this paper illustrates the application of SVM based classification of continuous wavelet features for fault diagnosis of monoblock centrifugal pump.

The rest of the paper is organized as follows. In Section 2, experimental setup and experimental procedure are described followed by feature extraction from the time domain signal is presented in Section 3. Then SVM classifier and the results of the experiment are discussed in Sections 4 and 5 respectively. Finally, conclusions are presented in Section 6 followed by references.

2. Experimental studies

The different fault conditions considered for the study are bearing fault, impeller defect, bearing and impeller defect together and cavitation. The main focus of this study is on the application of SVM algorithm for fault diagnosis of monoblock centrifugal pump.

2.1. Experimental procedure

The vibration signals are measured from the monoblock centrifugal pump working under normal condition at a constant rotation speed of 2880 rpm. The vibration signal from accelerometer mounted on the pump inlet was taken. The sampling frequency was 24 kHz and sample length was 1024 for all conditions of the pump. 250 trials were taken for each monoblock centrifugal pump condition, and vibration signals were stored in the data files.

In the present study the following faults were simulated

(i) Cavitation
(ii) Bearing fault
(iii) Impeller fault
(iv) Bearing and Impeller fault together

2.2. Experimental result

The faults were introduced one at a time and the pump performance characteristic and vibration signals were taken. As a result of the experiment, the representative time domain plots are given in Fig. 2.

3. Feature extraction

The time domain signal can be used to perform fault diagnosis by analyzing vibration signals obtained from the experiment. Continuous Wavelet Transform (CWT) has been widely used and provides the physical characteristics of time–frequency domain data. Wavelet analysis of vibration signals yields different descriptive parameters. Fairly a wide set of parameters were selected as the basis for the study. A set of statistical parameters and histogram features have been extracted. From the pool of features the best ones were selected for classification. The wavelet transformations are explained below. In this paper, CWT of different versions of different wavelet families has been considered for different levels. The list of families considered for this study is given below:

1. Daubechies wavelet (db1,db2,db3,db4,db5,db6,db7,db8,db9,-db10).
2. Coiflet (coif1,coif2,coif3,coif4,coif5).
3. Bi-orthogonal wavelet (bior1.1,bior1.3,bior1.5,bior2.2,bior2.4,-bior2.6,bior2.8,bior3.5,bior3.7,bior3.9,bior4.4,bior5.5, bior6.8).
4. Reverse bi-orthogonal wavelet (rbio1.1,rbio1.3,rbio1.5,rbio2.2,-rbio2.4,rbio2.6,rbio2.8,rbio3.1,rbio3.3,rbio3.5,rbio3.7,rbio3.9,r-bio4.4,-rbio5.5,rbio6.8).
5. Symlets (sym2,sym3,sym4,sym5,sym6,sym7,sym8).
6. Meyer wavelet.
7. Morlet.
8. Gaussian wavelet (gaus1,gaus2,gaus3,gaus4,gaus5,gaus6,-gaus7,gaus8).

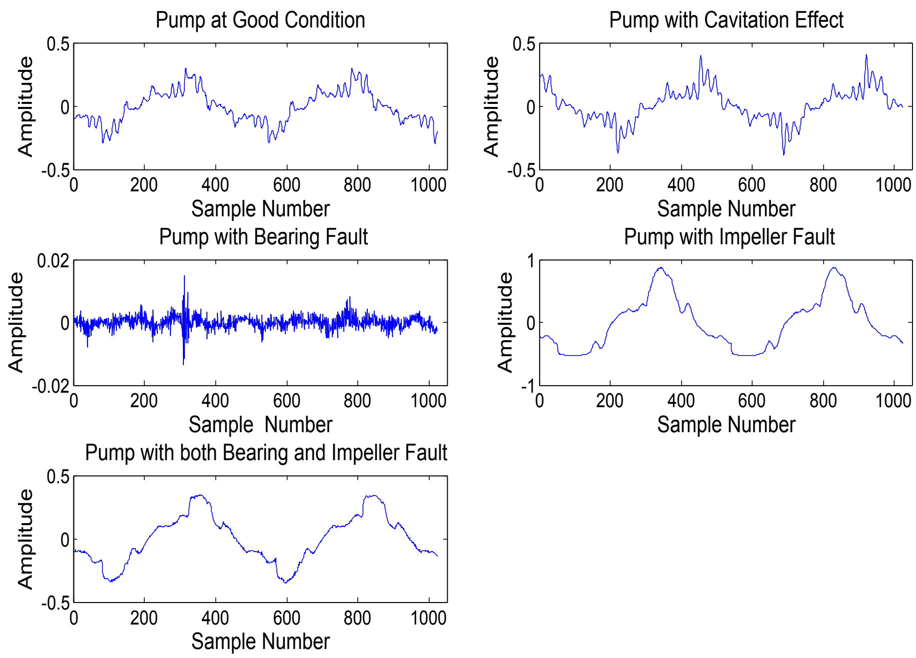

Fig. 2. Time domain plots of monoblock centrifugal pump.

3.1. Feature definition

3.1.1. Concept of continuous wavelet transforms

Wavelet transform is transform for time–frequency analysis. Wavelet means 'small wave'. Short duration finite energy functions can be called as wavelets. Wavelet transforms a signal under investigation like a vibration signal into another representation which presents the signal in a more useful form. Wavelet transform is a time-scale representation of a signal. Wavelet theory has been developed and applied widely in the recent years. The step by step procedure for the calculation of continuous wavelet transform features can be better understood in the form a flow chart. Fig. 3 illustrates the flow of the process.

A continuous wavelet transform is defined as

$$W(a,b) = \int_t f(t) \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) dt \quad (1)$$

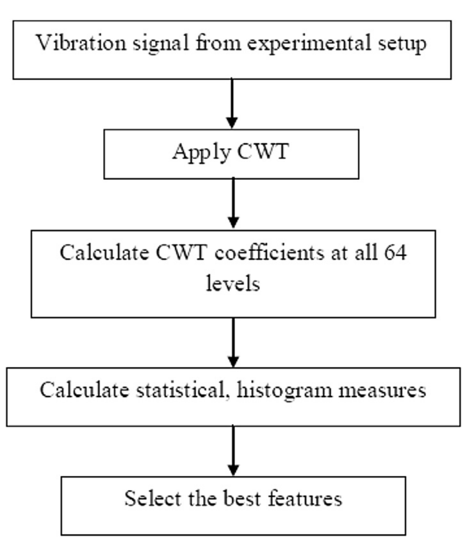

Fig. 3. Flow chart for feature extraction.

where,

ψ – wavelet function,
a – scaling parameter,
b – location parameter.

Assume that $f(t)$ is a complex valued function on $\Re$ which represent some signal (think of t as time).

The Fourier transform

$$\hat{f}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-\mathrm{it}.\omega} dt \quad (2)$$

is used to decompose f into its frequency components. The inversion formula

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{\mathrm{it}.\omega} d\omega \quad (3)$$

can be interpreted as writing f as a superposition of time-harmonic waves $e^{\mathrm{it}.\omega}$. If $\hat{f}$ is large near some frequency, then f has a large component that is periodic with that frequency. This approach works well for analyzing signals that are produced by some periodic process. Consider now two methods that attempt to provide information on both time and frequency: the *Windowed Fourier Transform* (WFT), also called *Short Time Fourier Transform* (STFT), and the *Continuous Wavelet Transform* (CWT). Both of them map a function of one variable (time) into a function of two variables (time and frequency). A large value of the transform near time t, frequency ω is interpreted as: the signal f contains a large component with frequency ω near time t. Similar to FFT, a fast wavelet transform is also available for computation work. One main advantage of wavelet transform is for signals with long duration low frequencies and short duration high frequencies. It has ability to produce a high frequency resolution at low frequencies and a high time resolution at high frequencies. Another advantage of wavelet transform is its ability to reduce noise in raw signals. Fig. 4

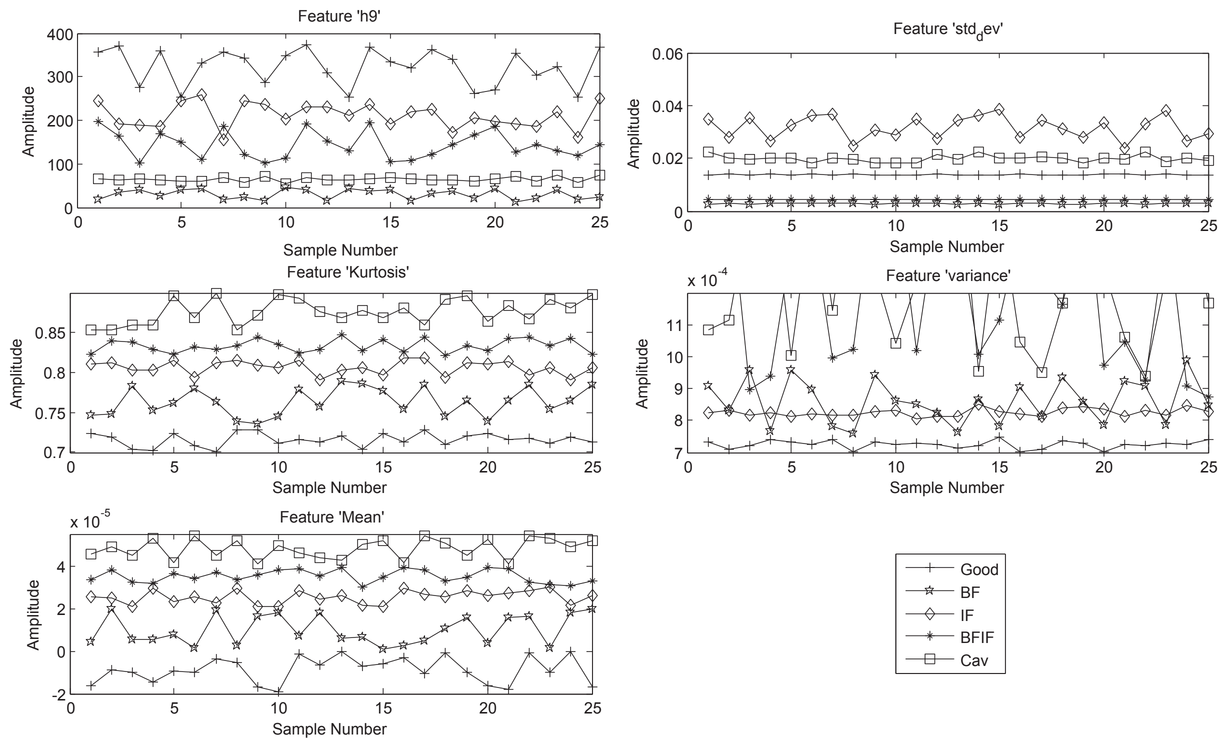

Fig. 4. Scatter plot of selected features.

shows the scatter plot of selected features against classification of faults.

4. SVM classifier

Data mining techniques are being increasingly used in many modern organisations to retrieve valuable knowledge structures from databases, including vibration data. An important knowledge structure that can result from data mining activities is SVM. As depicted in Fig. 5.

Consider the problem of classifying m points in the n-dimensional real space R^n, represented by the $m \times n$ matrix A and $m \times m$ diagonal matrix D with plus ones or minus ones along its diagonal according to membership of each point O_i in the class A+ or A−.

If, in the input space, data are not linearly separable, SV machines maps the data into some other dot product space called *feature space* **F** via a non-linear map $\phi: R^n \to \mathbf{F}$ and perform the above linear algorithm in **F**. This only requires the evaluation of dot products, $\phi(\mathbf{x}_i)^{\mathrm{T}}\phi(\mathbf{x})$. In literature, this function is called **kernel** and we denote $\mathbf{K}(\mathbf{x},\mathbf{y}) = \phi(\mathbf{x})^{\mathrm{T}}\phi(\mathbf{y})$.

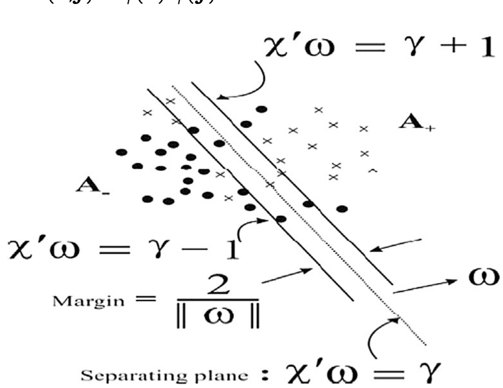

Fig. 5. SVM formulation with soft margin (boundary).

In the feature space, we find a linearly separating hyperplane $\mathbf{w}^{\mathrm{T}}\phi(\mathbf{x}) - \gamma = 0$; $\mathbf{w} \in \mathbf{F}$ such that the two classes are maximally separated.

Again **w** can be shown to be equal to $\sum_{i=1}^{k} \alpha_i \phi(\mathbf{x}_i)$ where k is the number of support vectors.

The decision function can now be written as :

$$f(\mathbf{x}) = \operatorname{sign}\left(\mathbf{w}^{\mathrm{T}}\phi(\mathbf{x}) - \gamma\right) = \operatorname{sign}\left(\sum_{i=1}^{k} \alpha_i \left(\phi(\mathbf{x}_i)^{\mathrm{T}}\phi(\mathbf{x}) - \gamma\right)\right) \tag{4}$$

If **F** is high dimensional, $\phi(\mathbf{x}_i)^{\mathrm{T}}\phi(\mathbf{x})$ is very expensive to compute. However, there are simple forms of kernels that can be evaluated efficiently. One good example is the polynomial kernel where $\mathrm{k}(\mathbf{x},\mathbf{y}) = (\mathbf{x}^{\mathrm{T}}\mathbf{y})^d$.

For $d = 2$ and $\mathbf{x}, \mathbf{y} \in R^2$, we have

$$\begin{aligned}\left(x^{\mathrm{T}}y\right)^2 &= \left[(x_1, x_2)^{\mathrm{T}}\left(y_1, y_2\right)\right]^2 = [x_1y_1 + x_2y_2]^2 \\ &= (x_1y_1)^2 + 2x_1y_1x_2y_2 + (x_2y_2)^2 \\ &= \left[\left(x_1^2, x_2^2, \sqrt{2}x_1x_2\right)^{\mathrm{T}}\left(y_1^2, y_2^2, \sqrt{2}y_1y_2\right)\right] = \phi(\mathbf{x}_i)^{\mathrm{T}}\phi(\mathbf{y})\end{aligned}$$

here ϕ maps (x_1, x_2) to $[(x_1^2, x_2^2, \sqrt{2}x_1x_2)]$.

This is so-called Kernel *trick*. We do not actually map the data into feature space for any computation. Computation is done in the input space itself. In the example, $(\mathbf{x}^{\mathrm{T}}\mathbf{y})^2$ is equivalent to $\phi(\mathbf{x}_i)^{\mathrm{T}}\phi(\mathbf{y})$. The quantity $(\mathbf{x}^{\mathrm{T}}\mathbf{y})$ is a scalar and after computing that we just square it to get the scalar quantity $\phi(\mathbf{x}_i)^{\mathrm{T}}\phi(\mathbf{y})$. The problem of obtaining **w** as a linear combination of subset of training samples (in feature space F) called *support vectors* are obtained by formulating and solving the problem as a quadratic programming problem. The training algorithm uses Sequential Minimal Optimization (SMO) technique. Once we obtain the support vectors using the

SMO algorithm [9], the classification of new data point **x** requires only the computation of sign$(\sum_{i=1}^{k}\alpha_i(\phi(\mathbf{x}_i)^{\mathrm{T}}\phi(\mathbf{x})-\gamma))$, where **i** is the index of support vectors.

5. Results and discussion

The experimental studies have been carried out for good condition and various fault conditions of the pump. One should understand that, for all faults considered in the study the performance of the pump is adversely affected. Hence, the study is important with these faults. In the present study SVM algorithm is used as explained in the previous Section 4. The input of SVM algorithm is set of wavelet features which are extracted from the vibration signals as explained in Section 3.

From the selected features the classification has been carried out. In order to have higher prediction accuracy and to avoid over fitting of data, a set of experiments were carried out to design the classifier and the results are discussed below. The results (classification accuracies) obtained from SVM classifier using wavelet features of various wavelet families shall be better explained in this manner. As a first step, the classification accuracy is found for different versions of the wavelet family. In the similar fashion, the efficiencies of different versions of all the mentioned wavelet families are computed and plotted as histogram charts.

Now, from the Figs. 6–11, best version of different families were picked up from each of the chart and compared among those best versions of different wavelet families and the overall best wavelet family and the best version of that family were found. In this manner, the versions bior3.7 (99.76), coif3 (99.76%), db8 (99.84%), gaus6 (99.68), rbio1.5 (99.68%), sym7 (99.68%), dmey (97.68) and mor (97.68%) have been picked up as the best from the Figs. 4–9 respectively.

All the best versions of different wavelet families were compared. The overall best wavelet and the wavelet family were found and plotted as a histogram chart as shown in Fig. 12. From Fig. 12, one can clearly say that the best wavelet from the chart is db8 and the classification accuracy achieved is 99.84% which is very high among other families. The results of the wavelet db8 can be

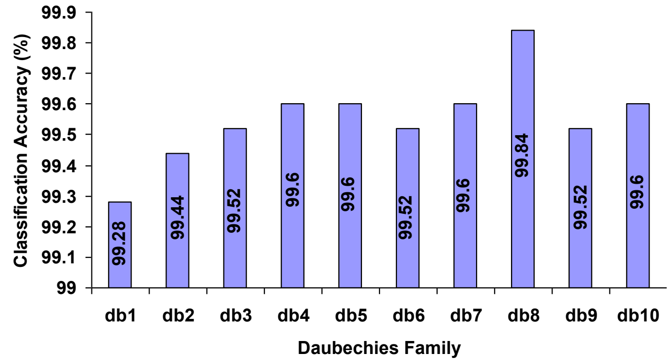

Fig. 8. Daubechies vs classification accuracy (%).

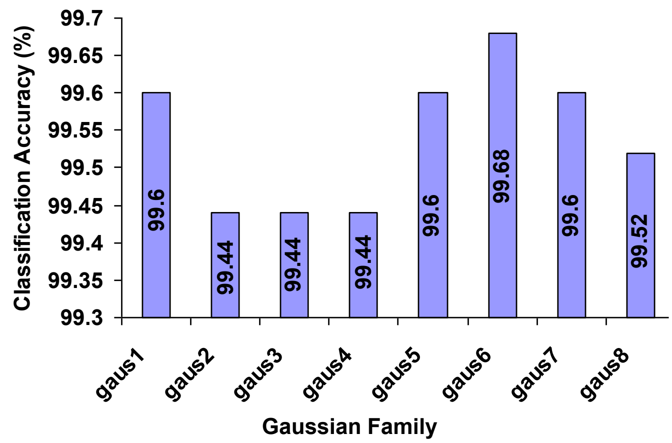

Fig. 9. Gaussian vs classification accuracy (%).

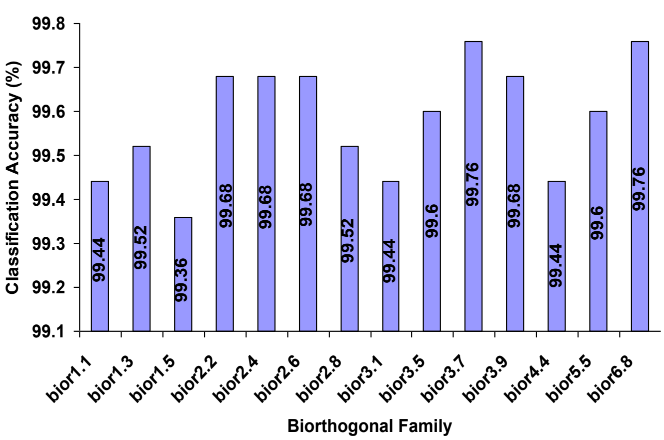

Fig. 6. Bi-orthogonal wavelet vs classification accuracy (%).

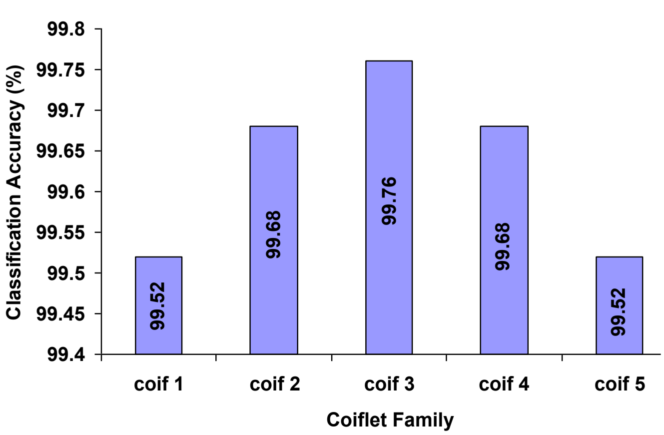

Fig. 7. Coiflet vs classification accuracy (%).

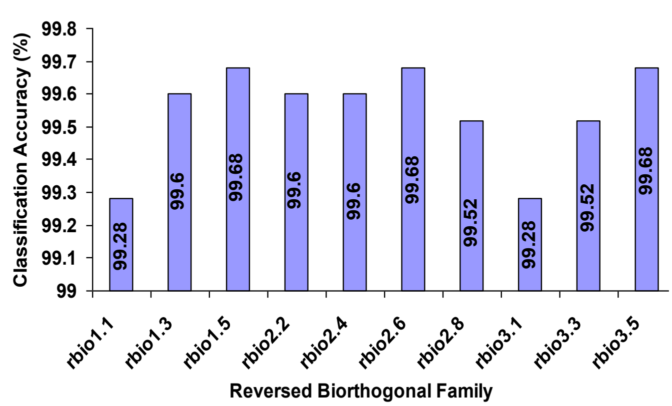

Fig. 10. Reversed bi-orthogonal wavelet vs classification accuracy (%).

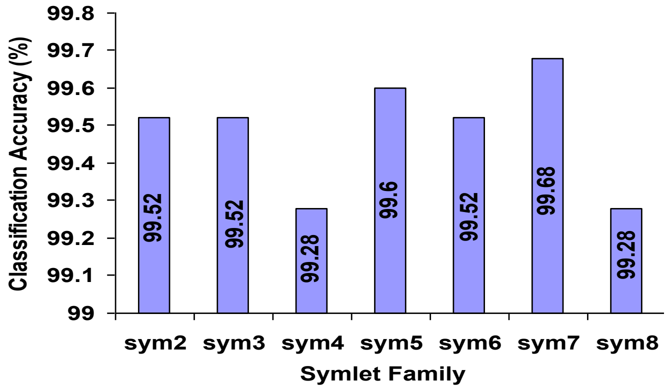

Fig. 11. Symlet vs classification accuracy (%).

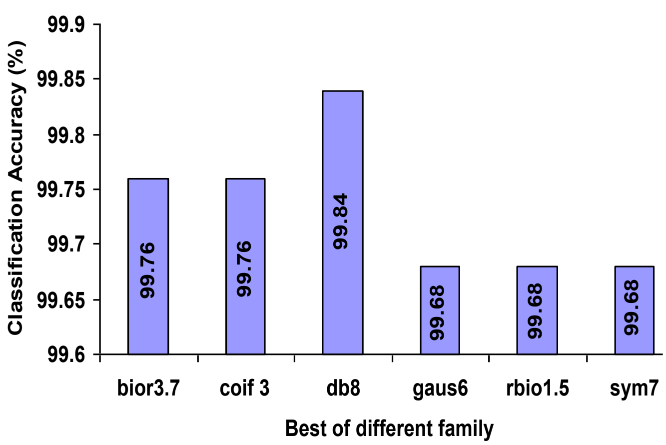

Fig. 12. Best of different family vs classification accuracy (%).

Table 1
Confusion matrix for db8 wavelet.

	Good	Cav	FI	FB	FBI
Good	250	0	0	0	0
Cav	0	250	0	0	0
FI	0	0	248	0	2
FB	0	0	0	250	0
FBI	0	0	0	0	250

illustrated in better way using the confusion matrix as shown in Table 1.

From the confusion matrix, one can understand that 250 samples were considered for each condition of the pump. All the diagonal elements of the confusion matrix represent the number of correctly classified data points and the non-diagonal elements represent the incorrectly classified data points. In this fashion the classification accuracies are found and compared for various types of wavelets of different families. In this case, all the good condition data points have been correctly classified and the same is the case with bearing fault data points and fault with both bearing and impeller. However, there were two misclassifications in pump with faulty impeller and they were classified as pump with faulty bearing and impeller conditions. Hence, efficiency was calculated to be 99.84%. The results obtained are specific to this particular dataset. Classification accuracy of 99.84% does not assure similar performance for all feature datasets. However one can expect classification accuracy close to 100%. In general the classification accuracy is very high. Hence the db8 wavelet is very much suited for fault diagnosis of centrifugal pumps.

6. Conclusion

This paper deals with vibration based fault diagnosis of monoblock centrifugal pump. Five classical states viz., normal, cavitation, bearing fault, impeller fault, impeller and bearing fault together, are simulated on monoblock centrifugal pump. Set of features have been extracted using different wavelets and classified using SVM algorithm. From the results and discussion as discussed above one can confidently say that feature extraction using wavelets and SVM algorithm for classification are good candidates for practical applications of fault diagnosis of monoblock centrifugal pump.

References

[1] H.Q. Wang, P. Chen, Mie, Fault diagnosis of centrifugal pump using symptom parameters in frequency domain, CGIR J. 9 (2007) 1–14.
[2] V. Muralidharan, V. Sugumaran, A comparative study of Naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis, Appl. Soft Comput. 12 (8) (2012) 2023–2029.
[3] V. Muralidharan, V. Sugumaran, N.R. Sakthivel, Wavelet decomposition and support vector machine for fault diagnosis of monoblock centrifugal pump, Int. J. Data Anal. Tech. Strat. 3 (2) (2011) 159–177.
[4] Kemal Polat, Salih Gunes, A novel hybrid intelligent method based on C4.5 decision tree classifier and one-against-all approach for multi-class classification problems, Expert Syst. Appl. 36 (2009) 1587–1592.
[5] V. Muralidharan, V. Sugumaran, Selection of discrete wavelets for fault diagnosis of monoblock centrifugal pump using the J48 algorithm, Appl. Artif. Intell. 27 (2013) 1–19.
[6] Fansen Konga, Ruheng Chen, A combined method for triplex pump fault diagnosis based on wavelet transform, fuzzy logic and neuro-networks, Mech. Syst. Signal Process. 18 (2004) 161–168.
[7] Jiangping Wang, Hongtao Hu, Vibration-based fault diagnosis of pump using fuzzy technique, Measurement 39 (2006) 176–185.
[8] Javier Sanz, Ricardo Perera, Consuelo Huerta, Fault diagnosis of rotating machinery based on auto associative neural networks and wavelet transforms, J. Sound Vib. 302 (2007) 981–999.
[9] V. Muralidharan, V. Sugumaran, Feature extraction using wavelets and classification through decision tree algorithm for fault diagnosis of mono-block centrifugal pump, Measurement 46 (2013) 353–359.
[10] Changqing Shen, Dong Wang, Fanrang Kong, Peter W. Tse, Fault diagnosis of rotating machinery based on the statistical parameters of wavelet packet paving and a generic support vector regressive classifier, Measurement 46 (2013) 1551–1564.
[11] Feng Li, Baoping Tang, Rongsong Yang, A novel fault diagnosis model for gearbox based on wavelet support vector machine with immune genetic algorithm, Measurement 46 (2013) 220–232.
[12] Zhixiong Li, Xinping Yan, Zhe Tian, Chengqing Yuan, Zhongxiao Peng, Li Li, Blind vibration component separation and nonlinear feature extraction applied to the nonstationary vibration signals for the gearbox multi-fault diagnosis, Measurement 46 (2013) 259–271.
[13] Jin Yuan, Xuemei Liu, Semi-supervised learning and condition fusion for fault diagnosis, Mech. Syst. Signal Process. 38 (2013) 615–627.
[14] Andre Lemos, Walmir Caminhas, Fernando Gomide, Adaptive fault detection and diagnosis using an evolving fuzzy classifier, Inf. Sci. 220 (2013) 64–85.