World Academy of Science, Engineering and Technology
International Journal of Mechanical and Mechatronics Engineering
Vol:2, No:7, 2008

A Comparative Study of SVM Classifiers and Artificial Neural Networks Application for Rolling Element Bearing Fault Diagnosis using Wavelet Transform Preprocessing

Commander Sunil Tyagi

International Science Index, Mechanical and Mechatronics Engineering Vol:2, No:7, 2008 waset.org/Publication/9372

***Abstract*—**Effectiveness of Artificial Neural Networks (ANN) and Support Vector Machines (SVM) classifiers for fault diagnosis of rolling element bearings are presented in this paper. The characteristic features of vibration signals of rotating driveline that was run in its normal condition and with faults introduced were used as input to ANN and SVM classifiers. Simple statistical features such as standard deviation, skewness, kurtosis etc. of the time-domain vibration signal segments along with peaks of the signal and peak of power spectral density (PSD) are used as features to input the ANN and SVM classifier. The effect of preprocessing of the vibration signal by Discreet Wavelet Transform (DWT) prior to feature extraction is also studied. It is shown from the experimental results that the performance of SVM classifier in identification of bearing condition is better then ANN and pre-processing of vibration signal by DWT enhances the effectiveness of both ANN and SVM classifier

***Keywords*—**ANN, Artificial Intelligence, Fault Diagnosis, Pattern Recognition, Rolling Element Bearing, SVM. Wavelet Transform

I. Introduction

ROLLING element bearings (REB) are most common element used in rotating machinery and their failure is the foremost cause of down time in plant machinery.

Most common REB defects are cracks or pits located at outer race, inner race and on the rolling element. These defects generate a series of impacts as rolling element passes over the defect due to the metal to metal contact. The resultant vibration is characterised by sharp peaks. It is difficult to identify the defect frequency in spectrum as these impact vibrations distributes their energy over wide range of frequencies; the bearing's defect frequency contains low energy [1] and hence can be easily masked by noise and other low frequency effects.

To overcome this problem, both time and frequency domain methods have been developed [2].Time domain methods usually involve indices that are sensitive to impulsive oscillations, such as peak level, rms value, crest factor analysis, kurtosis and shock pulse counting [3]–[6]. Since it is difficult to identify bearing defect in direct spectrum [7] many spectral techniques have been developed over the years for bearing fault diagnosis, such as Adaptive Noise cancellation [8], High Frequency Resonant Technique or Envelope Detection [9] and Wavelet Transform [10]. More recently various researchers have applied ANN [11]-[14] and SVM classifiers [23]-[26] to machinery fault diagnosis.

In the present work, a comparative study is presented on effectiveness of ANN and SVMs for bearing fault diagnostics using time-domain as well as frequency spectrum features. The vibration signals obtained from bearing in normal condition and bearings induced with faults are subjected to direct and simple processing for extraction of features that are subsequently used as inputs to the ANN and SVM classifier for diagnosing the bearing condition of a rotating machine. In the present approach, sets of normalized features are used so that even if the signals change in magnitude due to the change in speed or quality of sensor mounting, the diagnostic results are unaffected as long as the signal patterns remain unchanged. The features are obtained from the segments of the measured vibration signals instead of single values like crest factor, kurtosis and peaks for the undivided signals [11], [12].The effects of different types of bearing faults, different features and preprocessing by DWT are studied. A procedure is presented to correctly categorise the bearing conditions. The procedure is illustrated using the vibration data of a rotating shaft-line with normal and defective bearings.

II. Artificial Neural Networks

Artificial neural networks (ANN) are simplified artificial models based on the biological learning process of the human brain [15]. ANN has been very extensive in recent years such as in prognosis, classification, function approximation, control filter, pattern recognition etc [16]. Various researchers have used ANN for machinery fault detection. Application of ANN to preprocess, compress and classify vibration spectrum for bearing faults have been demonstrated by Alguindigue [12]. Youshang [11] presented a method for classification of bearing faults by ANN that was fed from DWT preprocessed signals. Wu and Liu [13] used ANN along with DWT to

Commander Sunil Tyagi is with the Directorate of Naval Architecture Integrated Headquarters of MoD (Navy), New Delhi, Pin 110011 India (phone: +91 11 23074363; fax: +91 11 23011134. 23010126; e-mail: tyagisunil@ hotmail.com).

International Science Index, Mechanical and Mechatronics Engineering Vol:2, No:7, 2008 waset.org/Publication/9372

investigate various faults in an internal combustion engine. Rajakarunakaran [14] used two different ANN techniques backpropagation algorithm and adaptive resonance network for fault detection of a centrifugal pump.

An ANN consists of a number of interconnected artificial processing neurons called nodes, connected together in layers forming a network. A typical ANN is schematically illustrated in Fig. 1. This is known as a two-layered ANN as the input layer performs no calculations. The number of nodes within the input and output layers are dictated by the nature of the problem to be solved and the number of input and output variables needed to define the problem. The number of hidden layers and the nodes within each hidden layer is usually a trial and error process.

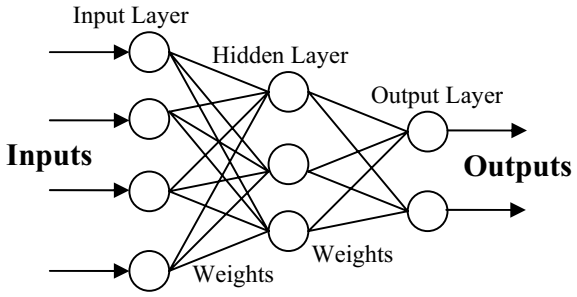

Fig. 1 Artificial Neural Network

As illustrated in Fig.2 each node in a layer (except the ones in the input layer) provides a threshold of a single value by summing up their input value p_i with the corresponding weight value w_i. Then the neuron's net input value n is formed by adding up this weighted value (sum), with the bias term b. The bias is added to shift the sum relative to the origin. The net input value then goes into transfer function f, which produces the neuron output a.

$$a = f\left(\sum_{i=1}^{r} w_i \cdot p_i + b\right) \tag{1}$$

The transfer function f that transforms the weighted inputs into the output a is usually a non linear function. The sigmoid (S-shaped) or logistic function is the most commonly used transfer function which restricts the nodes output between 0 and 1.

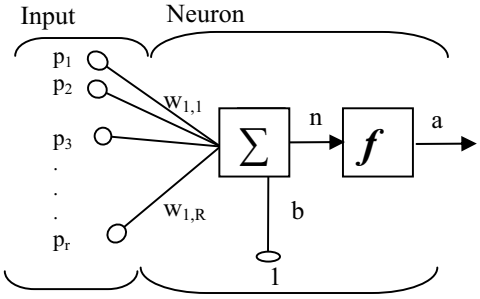

Fig. 2 A neuron

A The procedure most commonly used to train an ANN is a method known as backpropagation [17]. This is a supervised method of learning mainly used to train multilayer neural networks. In supervised learning, a set of inputs are applied to the network, then the resultant outputs produced by the network are compared with that of the desired ones. If the network is provided with following set of examples for proper behavior:

$$\{p_1,t_1\}\ ,\ \{p_2,t_2\}\ ,\ \ldots\ ,\ \{p_Q,t_Q\}, \tag{2}$$

where p_Q is an input to network and t_Q is corresponding target.

The normalized mean square error (MSE) is calculated and propagated backwards via the network. Back propagation network (BPN) uses it to adjust the value of the weights on the neural connection in the multiple layers. This process is repeated until the MSE is reduced to an acceptably low value, which would be suitable to classify the test set correctly. The mean square error function $F(x)$ at iteration k is given by:

$$F(x) = \left[\left(t_k - a_k\right)^2\right] \tag{3}$$

BPN uses steepest descent method to adjust the weights and biases. The adjusted weights and biases of mth layer at iteration k are estimated by:

$$w_{i,j}^m(k+1) = w_{i,j}^m(k) - \alpha \frac{\partial F}{\partial w_{i,j}^m} \tag{4}$$

$$b_i^m(k+1) = b_i^m(k) - \alpha \frac{\partial F}{\partial b_i^m} \tag{5}$$

where α is learning rate and $w_{i,j}$ represents weights of connection between neuron i and neuron j.

After the ANN is successfully trained, it should be ready to test data not seen previously. Various algorithms are available to implement backpropogation network most common amongst them is Levenberg-Marquardt algorithm [18] which has been used in this paper.

III. Support Vector Machines (SVMs)

ANNs have proven good classifiers but they require large number of samples for training, which is not always true in practice [19]. Support vector machines (SVMs) are based on statistical learning theory and they specialise for a smaller sample number. SVMs have better generalisation than ANNs and guarantee the local and global optimal solution similar to that obtained by ANN [20]. In recent years, SVMs have been found to be remarkably effective in many real-world applications [21],[22]. As it is hard to obtain sufficient fault samples in practice, SVMs have been applied for machinery fault diagnosis by various researchers in recent times [23]-[26]. Yang [26] used intrinsic mode function envelop spectrum as input to SVMs for classification of bearing faults

International Science Index, Mechanical and Mechatronics Engineering Vol:2, No:7, 2008 waset.org/Publication/9372

and Hu [25] used improved wavelet packets and SVMs for the bearing fault detection.

SVM is developed from the optimal separation plane under linearly separable condition. Its basic principle can be illustrated in two-dimensional way as Fig.3 [27]. Fig.3 shows the classification of a series of points for two different classes of data, class ***A*** (circles) and class ***B*** (pentacles). The SVM tries to place a linear boundary ***H*** between the two classes and orients it in such way that the margin is maximized, namely, the distance between the boundary and the nearest data point in each class is maximal. The nearest data points are used to define the margin and are known as support vectors.

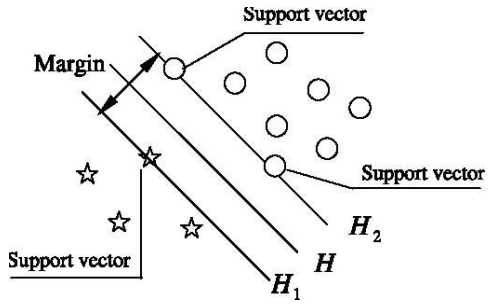

Fig. 3 Classification of data by SVM

Suppose there is a given training sample set G={(xi,yi), i=1...*l*}, each sample $\mathbf{x_i} \in R^d$ belongs to a class by $y \in \{+1,-1\}$.

The boundary can be expressed as follows:

$$\omega \cdot x + b = 0 \tag{6}$$

where **ω** is a weight vector and **b** is a bias. So the following decision function can be used to classify any data point in either class **A** or **B**:

$$f(x) = \operatorname{sgn}(\omega \cdot x + b) \tag{7}$$

The optimal hyperplane separating the data can be obtained as a solution to the following constrained optimization problem:

$$\text{Minimise } \frac{1}{2}\|\omega\|^2 \tag{8}$$

$$\text{Subjet to } y_i\left[(\omega \cdot x + b)\right] - 1 \geq 0, \quad i = 1,....l \tag{9}$$

Introducing the Lagrange multipliers $\alpha_i \geq 0,$ the optimization problem can be rewritten as:

$$\text{Maximise } L(\omega, b, \alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \tag{10}$$

$$\text{Subject to } \sum_{i=1}^{l} \alpha_i y_i = 0 \tag{11}$$

The decision function can be obtained as follows

$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{l} \alpha_i y_i (x_i \cdot x) + b\right) \tag{12}$$

If the linear boundary in the input space **s** is not enough to separate into two classes properly, it is possible to create a hyperplane that allows linear separation in the higher dimension. In SVM, it is achieved by using a transformation **Φ(x)** that maps the data from input space to feature space. If a kernel function

$$K(x, y) = \Phi(x) \cdot \Phi(y) \tag{13}$$

is introduced to perform the transformation, the basic form of SVM can be obtained

$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{l} \alpha_i y_i K(x, x_i) + b\right) \tag{14}$$

Among the kernel functions in common use are linear functions, polynomials functions, radial basis functions multi layered perceptron and sigmoid functions.

IV. DWT and Multi-resolution analysis

A. Discreet Wavelet Transform

DWT have found wide applications in machinery fault diagnosis [29], [13] for their capability to treat the transient signals as it provides time and frequency representation together with multi resolution analysis [28]. Recently, wavelet transform has been applied for rolling element bearing fault diagnosis [10], [11], [30].

The wavelet transform is a tool that cuts up data, functions or operators into different frequency components, and then studies each component with solution matched to its scale. The use of wavelet transform is appropriate to analyze non-stationary signal since it gives the information about the signal both in frequency and time domains [31]. Let *x(t)* be the signal. The continuous wavelet transform (CWT) of *x(t)* is defined as:

$$W_\psi(\tau, s) = \int_{-\infty}^{+\infty} x(t) \cdot \psi^*_{\tau,s}(t)dt \tag{15}$$

Where $\psi^*_{\tau,s}(t)$ is conjugate of $\psi_{\tau,s}(t)$, that is the scaled and shifted version of the transforming function, called a "mother wavelet", which is defined as:

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-\tau}{s}\right) \tag{16}$$

International Science Index, Mechanical and Mechatronics Engineering Vol:2, No:7, 2008 waset.org/Publication/9372

The transformed signal is a function of τ and s, the translation and scale parameters. The mother wavelet is a prototype for generating the other wavelet (window) functions. The scale parameter performs scaling operation on the mother wavelet. Each scale represents a frequency band. The term translation corresponds to time information in the transform domain; it shifts the wavelet along the time axis to capture the time information contained in the signal.

The DWT is derived from discretization of $W_\psi(\tau, s)$ given by:

$$DWT(j,k) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{\infty} x(t)\psi\left(\frac{t-2^j k}{2^j}\right) dt \qquad (17)$$

An efficient way to implement this scheme was developed by Mallat [32]. The basic step of wavelet algorithm is illustrated in Fig.4.The DWT is performed by process of decomposition in which the discreet signal ***x*** is convolved with a low pass filter ***L*** and a high pass filter ***H***, resulting in two vectors ***A*** and ***D***. The vectors ***A*** and ***D*** are down sampled to obtain ***cA1*** called approximate coefficient and ***cD1*** called detail coefficient. In down sampling the odd indexed elements of filtered signals are omitted so that the numbers of coefficients produced in decomposition are equal to the number of elements in the discreet signal ***x(t)***.

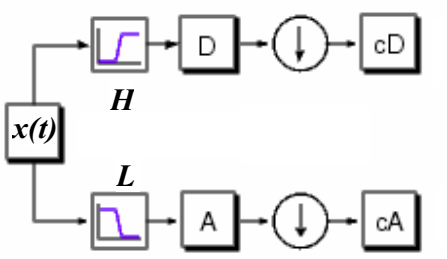

Fig. 4 Decomposition of wavelet Transform

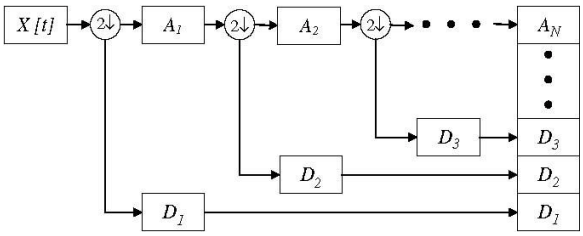

Fig. 5 Decomposition at different levels

B. Multi-Resolution Analysis

The decomposition process can be repeated using approximate coefficients ***cA*** to obtain DWT coefficients at different levels (scale) as per the desired resolution. The process is schematically depicted in Fig. 5.

V. Experimental Setup

The test rig shown in Fig. 6 was composed of a variable speed AC motor driving a shaft rotor assembly through flexible couplers; shafts were rested on two ball bearings. A rotor was used for balancing. The bearings under analysis (type MB 204) were placed at load end side for ease of replacement. The load on the system can be adjusted by a manually adjustable magnetic brake, which was driven via a belt drive. Vibration signals were acquired by accelerometer stud mounted on the bearing housing. The faults were artificially introduced to the bearings. The types of faults included a defective outer-race, a defective inner-race, and a defective roller. The shaft was made to rotate at 25 Hz and vibration signals were collected at sampling rate of 51.2 KSa/s. The numbers of samples collected were 102400 for duration of 2 s.

Following four signals were collected:

1. Bearing in normal condition,
2. Bearing with Outer Race fault,
3. Bearing with Inner Race fault, and
4. Bearing with Roller fault.

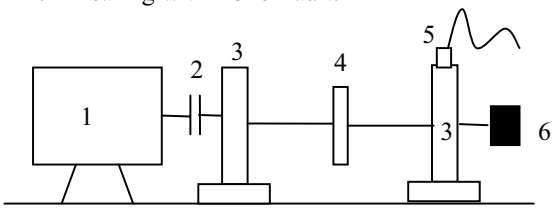

Fig. 6: 1. Variable speed motor 2. Flexible coupling 3. Bearing Housing 4. Rotor 5. Accelerometer 6. Magnetic break

VI. Features and Creation of Training/Test Vectors

A. Feature Selection

Each signal of 102400 samples was divided in 40 non overlapping bins of 2560 samples (y_i). Ten features were extracted from these 40 bins as follows:

Feature 1-5 - First five highest peaks
Feature 6 - Highest peak of power spectral density (PSD).
Feature 7 - Standard deviation σ.
Feature 8 - Skewness γ_3 (third central moment).
Feature 9 - Kurtosis γ_4 (fourth central moment).
Feature 10 - Sixth central moment γ_6.
The features 6 – 10 were extracted using:

$$\sigma = \sqrt{E\left\{(y_i-\mu)^2\right\}},\quad \gamma_3 = \frac{E\{(y_i-\mu)^3\}}{\sigma^3},\quad \gamma_4 = \frac{E\{(y_i-\mu)^4\}}{\sigma^4}$$

$$\gamma_6 = \frac{E\{(y_i-\mu)^6\}}{\sigma^6}$$, where $\mu = E\{y_i\}$ is the mean value and E is represents the expected value of the function.

International Science Index, Mechanical and Mechatronics Engineering Vol:2, No:7, 2008 waset.org/Publication/9372

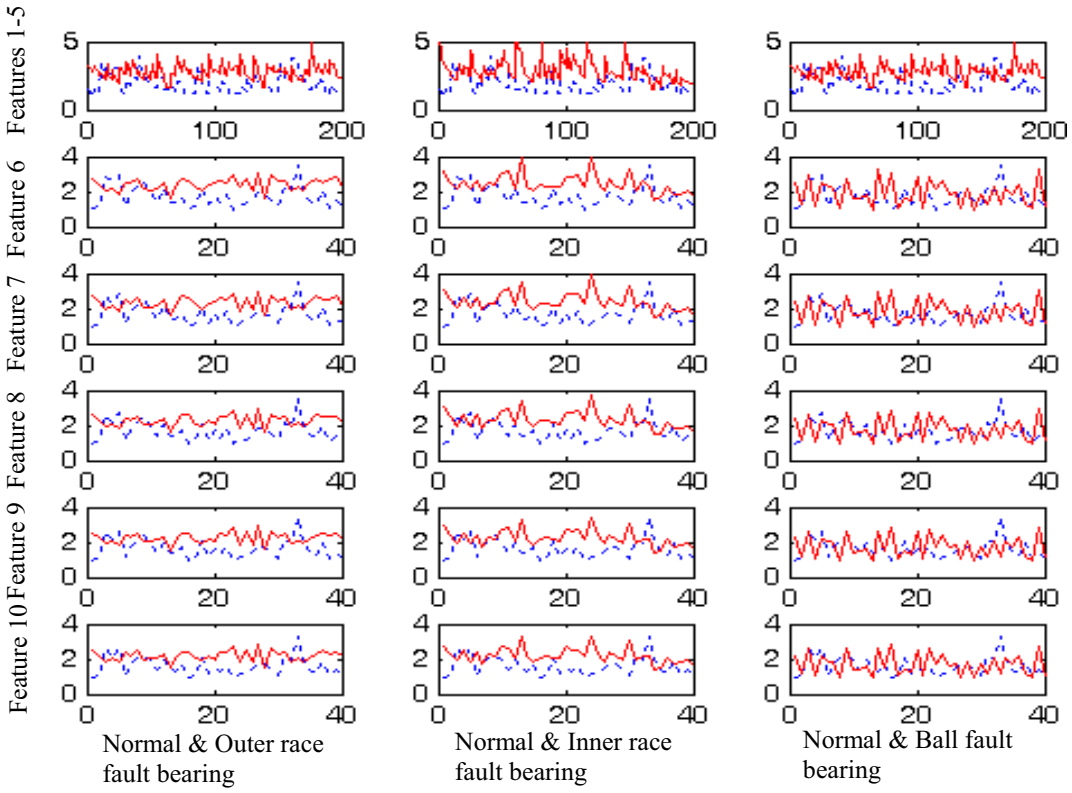

Fig. 7 Features of acquired Vibration signals —— defective, ---------- normal

Fig. 7 shows the plots of features extracted from the vibration signals. The features of defective bearings are plotted against that of normal bearing. The plots show good separation between the normal and the defective cases for all features justifying their selection.

These features extracted from vibration signals with or without the bearing fault were used for training the ANN and to input the SVM classifier for diagnosis of the bearing condition. The contribution of the features and the type of signals in the diagnosis of machine condition is discussed in section VII and VIII. Section IX presents the effect of preprocessing with DWT.

B. *Creation of training and test vectors*

The vibration signal 1–4 each with 102400 samples were divided into 40 bins each having length of 2560 samples. The lengths of bins were selected so that each would contain sufficient number (>5) of impacts caused by passing of the rolling element over the fault. Out of these bins 24 bins were used for training the ANN and SVM classifier. The remaining 16 bins which have not been seen by the ANN and SVM classifier were used for testing. The training sets were created by features extracted from defective signal bins and normal bearing signal bins alternately. Thus three sets of 48 training vectors outer race fault (ORF), inner race fault (IRF) and ball fault (BF), were created for outer race fault, inner race fault and ball fault respectively. Similarly three sets of 32 test vectors were also created. As there were 10 features, therefore a training matrix of 10X144 and test matrix of size 10X96 were created. The diagnostic capability of ANN and SVM classifiers for different faults were also studied by adding/omitting the training sets of respective signal. The numbers of features were also varied to measure their effect.

VII. Diagnosis of Bearing Condition using ANN

A. *Training of artificial neural network*

The neural network consisting of an input layer, three hidden layers and an output layer was used. The input layer has nodes representing the features extracted from the measured vibration signals. The number of neurons in the first hidden layer was varied from 10 to 30, the second one, from 5 to 10 and third from 2 to 10. The number of output nodes was varied between 1 and 2. The target values of two output nodes can have only binary values 1 or 0 representing normal and failed bearing. The ANN was trained using the MATLAB neural network toolbox using back propagation with Levenberg–Marquardt algorithm [18]. For training, a mean square error (MSE) of 10^{-5}, a minimum gradient of 10^{-10} and maximum iteration number (epoch) of 15000 were used. The training process would stop if any of these conditions were met. The initial weights and biases of the network were selected randomly.

The structure of the ANN giving best results was n:10:10:4:2 where the n represent the numbers of nodes in the input layer, the hidden layers had 10, 10 and 4 nodes and the

output layers had 2 nodes. In the training stage, the target value of the first output node for the normal bearing condition was set to 1 and that for the bearing having outer race, inner race and ball defect was set to 0.

B Effects of Bearing Defect Type

Table1 shows the results of training and testing the diagnostic capability of the ANN for different input vectors representing different type of bearing faults individually as well as in groups. All ten features were used to study the roles of the bearing defect type. Good training success (93.8 – 100%) was achieved in all cases studied. However the test success varied from 76 – 92.7%. The results of case 1 – 3 indicates that ANN is able to classify correctly for all type of bearing defects even if it is trained with features of only one type of defect (along with features of normal bearing)[1]. This indicates to similar nature of impact vibrations produced by different kind of bearing faults. Although 100% training success was achieved when ANN was input with ORF and BF signals; the test success was higher when trained with BF.

Results of case 4-6 clearly shows that that the contribution of ball fault signal is most significant for identification of bearing condition as both test and train success was lower (case 4) when features from BF signal was omitted. Case 7 gave best performance when all types of bearing defects were used to train the network.

C EFFECTS OF SIGNAL FEATURES

Table 2 shows the relative importance of signal features for identification of machine condition. For cases 18-19, all three input signals ORF, IRF and BF were used for training. The table presents results using all ten features, namely, the first five highest peaks, highest peak of PSD, standard deviation (σ), skewness (γ_3), kurtosis (γ_4) and sixth central moment (γ_6) either alone or in combination.. In cases 8–13, the ANN was trained with only one feature, the contribution of feature 10 i.e. sixth central moment (γ_6) was found most significant as it gave best test success of 94.8 % (91/96) the success with training set was also high 80.6%. In cases 9 and 12 i.e. when network was trained with only peak of PSD (feature 6) and kurtosis (γ_4) (feature 9) the performance goal could not be reached. However, these features were still retained as using them in combination with other features gave good results.

Low test results were obtained in case 15 and 18 when feature 6 and 9 were omitted. The use of central moments of order more than six did not have any significant effect on the diagnosis results. The third central moment γ_3 was found to be not a good feature as poor test success of only 40.6 % was obtained when the network was with only γ_3(case 11). Further, good test (97.9 %) and training success (98.6 %) was obtained in case 17 when γ_3 was omitted. In case 19 the training and test success were even better then case 7 which made use of all features. It is thus proposed that all features except feature 8 (γ_3) be used to train the ANN.

TABLE I
EFECT OF BEARING FAULT TYPE ON IDENTIFICATION OF MACHINE CONDITION

Case	Input signals	ANN		SVM	
		Training success	Test success	Training success	Test success
1	ORF	48/48 (100%)	74/96 (77.1%)	48/48 (100%)	77/96 (80.2%)
2	IRF	45/48 (93.8%)	74/96 (77.1%)	48/48 (100%)	83/96 (86.5%)
3	BF	45/48 (100%)	85/96 (88.5%)	48/48 (100%)	82/96 (85.4%)
4	ORF, IRF	93/96 (96.9%)	73/96 (76 %)	96/96 (100%)	84/96 (87.5%)
5	ORF, BF	96/96 (100%)	89/96 (92.7%)	96/96 (100%)	85/96 (88.5%)
6	IRF, BF	93/96 (96.9%)	78/96 (81.3%)	96/96 (100%)	88/96 (91.7%)
7	ORF, IRF,BF	141/144 (97.9%)	81/96 (84.4%)	144/144 (100 %)	90/96 (93.8%)

VIII. DIAGNOSIS OF BEARING CONDITION USING SVM CLASSIFIERS

The SVM classifiers was designed for same training and test vectors as used for ANN. Various kernel functions such as Linear, Quadratic, Multilayer Perceptron, Gaussian Radial

[1] Vectors ORF, BF & IRF contain features of defective and normal bearings.

TABLE II
EFFECT OF INPUT FEATURES ON IDENTIFICATION OF BEARING CONDITIONS

SVM	Test success (Max. 96)	68	54	76	50	59	55	87	75	80	95	89	89	90
	Training success (Max. 144)	129	85	123	94	106	106	141	144	144	144	144	144	144
ANN	Test success (Max. 96)	64	48	73	39	42	91	91	64	80	94	75	84	81
	Training success (Max. 144)	125	72*	122	124	123*	116	140	142	143	142	141	138	141
	Input features	1-5	6	7	8	9	10	6,7,8, 9,10	1-5,7, 8,9,10	1-5,6,8, 9,10	1-5,6,7, 9,10	1-5,6,7, 8,10	1-5,6,7, 8,9	1-5, 6,7, 8,9,10
	Case	8	9	10	11	12	13	14	15	16	17	18	19	7

International Science Index, Mechanical and Mechatronics Engineering Vol:2, No:7, 2008 waset.org/Publication/9372

International Science Index, Mechanical and Mechatronics Engineering Vol:2, No:7, 2008 waset.org/Publication/9372

Basis Function (RBF) and Polynomial kernels were used for all 19 cases.

The Linear, Quadratic and Multi layer perceptron kernels could not achieve convergence in many cases and training was stopped after maximum number of iterations (200000) was reached. The RBF and polynomial kernels achieved convergence in all 19 cases.

Fig. 8 presents the training, test and combined (test + train) success achieved by Polynomial and RBF kernel functions. The Fig. 8 presents the results by accumulating all 19 cases. The results are presented in percentages combining all 19 cases thus total 2304 training and 1824 test vectors were presented to the SVMs.

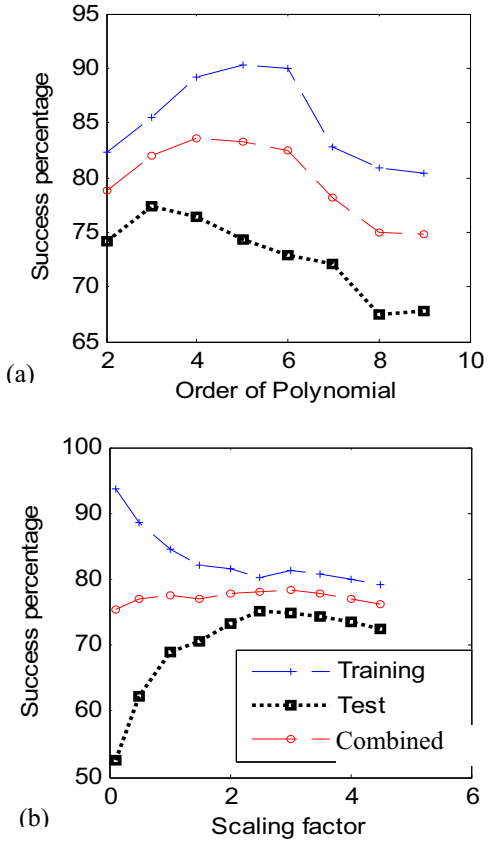

Fig. 8 SVM training (a) Polynomial kernel (b) RBF kernel

The order of polynomial for polynomial kernel was varied from 2 to 10 to find out the most optimum order. High combined (i.e. test plus training) success; more then 80 % was achieved for order of polynomial was set between 3 to 6. The test and train success increased as order of polynomial was increased up to 3. When the order was increased beyond 3 the SVM started to over fit; the training percentage continued to increase whereas the test success percentage fell. The 4th order of polynomial showed good balance of train and test success.

In case of training with RBF kernel the RBF scaling factor was varied from 0.1 to 5 to select the most optimum scaling factor. The test and train percentage increased as scaling factor was decreased from 5 to 2.5. The SVM starts to over fit as the scaling factor was decreased beyond 2.5. The cumulative percentages in case of RBF kernel were lower (<80%) as compared to the polynomial kernel. Polynomial kernel of the order of 4 was thus selected for the SVM classifier.

Three methods were tried to find the separating hyperplane namely Quadratic Programming, Least-Squares and Sequential Minimal Optimization method. The Sequential Minimal Optimization method was selected as it gave best results amongst the three choices.

The test and training results for 19 cases using SVM with 4th order polynomial kernel are presented at Table 1 and 2. It is evident from results that more accurate classification of bearing condition is achieved by using SVM classifiers as compared to ANN. Both the training as well as the test successes were higher in case of SVM classifier as compared to ANN for all cases except for case 13. In case 10 and 11 the train success was lower then that achieved by ANN but success with test vectors were higher.

Another significant aspect of using SVM classifiers is the speed of training. The time taken for SVM to train is far less in comparison to the time taken by the ANN. Maximum time taken by the SVM classifier to train was 4.6 sec (case 13). Amongst all cases the ANN achieved fastest training in case 1 where it took 12 sec to train the network. Thus even the slowest case of SVM was much faster then ANN. It is proposed that the SVM classifiers be used for bearing condition classification.

IX. Effects of Discreet Wavelet Transform

The effectiveness of pre-processing the acquired vibration signals by DWT is discussed in this section. Instead of using the raw signal as was the case in previous sections, the vibration signals were processed through DWT using Daubechies wavelet of order 4 (Db4) at level 6 to obtain the low frequency approximate at level six (A6) and the high frequency detail signals at level 1 to 6 (D1-D6). Frequency range of details were in the descending order, i.e. D1 had highest frequency content (12-25.6 kHz), and D6 had the lowest frequency content (0.3–1.2 kHz) whereas frequency content of D2 ranged from 4.6 – 10 kHz. The test and train vectors were created out of these details (D1 - D6) instead the raw signal. The selection of features and creation of training/test vectors were done as per section VI.

Fig. 9 shows the results of using various details (D1 – D6) by the SVM classifier. Total 2304 training and 1824 test vectors created (by accumulating all 19 cases) for each detail were presented to the SVM classifier. Fig. 9 presents the training, test and combined (training + test) success percentages achieved. High test and training success was achieved when details D1 and D2 were used, indicating the influence of the bearing defect in high frequency range > 5 kHz [7]. The performance of components D3 – D6 outside this frequency range were not very satisfactory. Best results were obtained when D1 was used to extract features, in this case the cumulative (case 1 – 19) training success was 94.3% and test success was 80 %.

The Fig. 10 presents (case wise) the combined (training plus test) success obtained by ANN when the network was

International Science Index, Mechanical and Mechatronics Engineering Vol:2, No:7, 2008 waset.org/Publication/9372

input with raw signal and with signals pre-processed with DWT (using level 1 details D1).

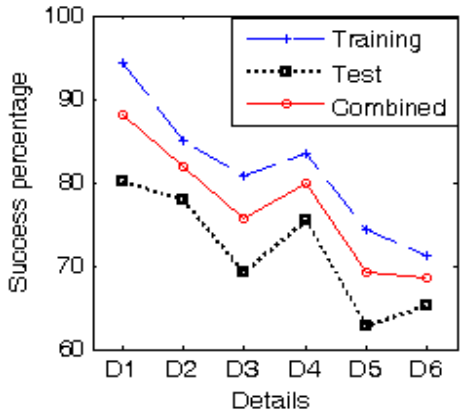

Fig. 9 Effect of DWT details at various levels

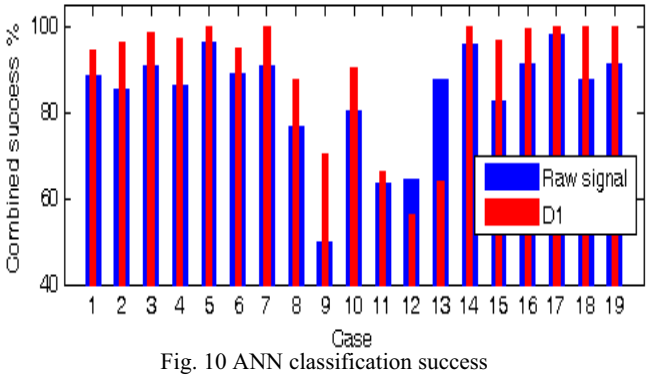

Fig. 10 ANN classification success

Pre-processing with DWT has improved the success (combined) of classifying the bearing condition in 17 cases out of total 19 cases. The maximum increase was in case 9 where combined success increased from 120 (50%) to 170(70.8%). By pre-processing with DWT the success reduced in only two case i.e. case 12 and case 13.

Pre-processing with DWT also improves the classification of bearing condition by the SVM classifiers. Fig 11 presents the combined (training + test) success achieved by the SVM classifier when input with features extracted from raw signal and from pre-processed signal (D1). The combined success increased in 14 out of 19 cases. The maximum increase was in case 9, whereas four cases i.e.12 – 14, 17 and 18 showed decrease.

Another significant effect of pre-processing the vibration signal with DWT was on the number of iterations (epochs) performed by the ANN and SVM classifier to train. Table 3 presents the number of iterations performed by ANN and SVM classifier to train when they were input with features extracted from raw signal and with signal pre-processed with DWT. When ANN was input with DWT processed signal the number of iterations required for training had reduced considerably for all cases except for case 8 and cases 12, 13 (where the training had stopped at 15000 epochs without reaching the performance goal).

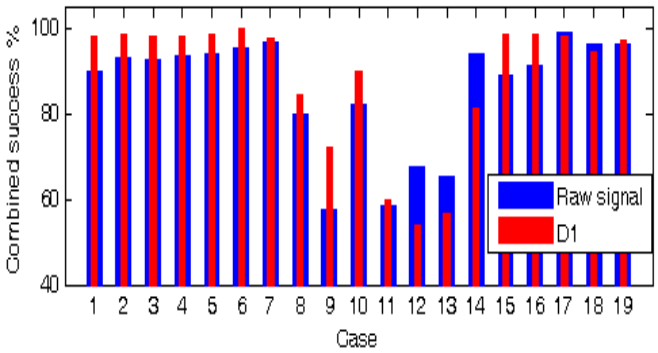

Fig. 11 SVM classification success

There has been more then 50% reduction in six cases and maximum reduction was in case 10 where the number of epochs required for training reduced by about 90 %.

Similarly when SVM was fed with features of signal preprocessed with DWT there is a reduction in number of iterations performed in training. The reduction was observed in 13 out of 19 cases. The number of training iterations had increased in cases 14 – 19 where the SVM classifier was fed with one features less. However, when all ten features were fed (case 7) the number of iterations reduced by about 82 %.

X. Conclusion

A method is presented to identify bearing condition by using simple features such as five highest peaks and statistical central moments of time domain vibration signal together with peak of Power Spectral Density. It is shown that using these simple features the bearing condition can be correctly classified with high accuracy with the help of ANN or SVM classifiers. 19 different cases were created to test the efficacy of ANN and SVM classifier in different conditions and it was found that SVM classifier performs better then ANN in almost all cases. Preprocessing with DWT improves the performance of both ANN and SVM classifiers. The test and train success increased in most cases when features extracted from details at level 1 (D1) were used to train ANN and SVM classifier. The DWT preprocessing also significantly lowers the number of iterations (epochs) required to train the ANN and SVM classifiers.

In practice it is difficult to obtain vibration signatures arising out of all kinds of bearing faults such as outer race fault, inner race fault or ball fault. In proposed method the vibration signals from any one type of bearing fault is sufficient to diagnose the bearing condition that may have other type of defect. It is perhaps because the proposed method does not attempt to make use of bearing defect frequency or time domain features; it focuses upon the peaky nature of impact vibrations by using highest peaks and statistical features such as central moments.

The present procedure is used to classify the status of the machine in the form of normal or faulty bearings. There is a scope for its extension to identify fault types and severity

levels. Since the SVM classifier training is quite fast, training and test may be done on-line. These issues are subjects for further study.

TABLE III NUMBER OF ITERATIONS PERFORMED DURING TRAINING

	ANN		SVM	
Case	**Raw Signal**	**D1**	**Raw Signal**	**D1**
1	12	9	128	58
2	18	14	365	88
3	33	16	789	204
4	27	13	532	71
5	32	20	647	113
6	27	25	506	159
7	32	27	550	101
8	106	152	6555	4433
9	15000[a]	3578	2614	158
10	2275	154	64	133
11	6760	4675	494	108
12	15000 [a]	15000 [a]	1567	142
13	11930	15000 [a]	6668	146
14	56	26	857	23536
15	64	32	6523	17429
16	37	32	821	49833
17	45	27	1154	36153
18	40	42	1023	26232
19	44	24	537	51587

[a] Training was terminated at maximum number of epochs

ACKNOWLEDGMENT

The work was supported by INS Shivaji, Lonavla, the Indian Navy's training and research establishment.

REFERENCES

[1] J. Pineyro, A. Klempnow and V. Lescano, Effectiveness of new spectral techniques in the anomaly detection of rolling element bearings," J of Alloys and Compounds,310, 2000, p.276-279.
[2] N. Tandon and A. Choudhury, "A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings," Tribology International, 32, 1999, p.469-80.
[3] T. Miyachi and K. Seki, "An investigation of early detection of defects in ball bearings using vibration monitoring- practical limit of detectability," Proceedings of the International Conference on Rotodynamics, JSME-IFToMM, Tokyo, 14-17 September,1986,p.403-8.
[4] D. Dyer and R.M. Stewart, "Detection of rolling element bearing damage by statistical vibration analysis," Trans ASME, J Mech Design, 100(2), 1978, p.229-35.
[5] A.A. Rush, "Kurtosis – A crystal ball for maintenance engineers," Iron and Steel Int, February 1979, p. 23-27.
[6] D.E. Butler, "The shock pulse method for the detection of damaged rolling bearings," NDT Int 1973, p.92-95.
[7] N. Tandon and BC.Nakra, "Detection of defects in rolling element bearings by vibration monitoring," J Instn Engrs (India) — Mech Eng Div 73,1993, p.271–82.
[8] G.K. Chaturvedi and DW. Thomas, "Bearing fault detection using adaptive noise canceling," ASME Paper 81-DET-7, New York, ASME, 1981, p.10.
[9] J. Courrech, "Envelop Analysis for effective rolling element bearing fault detection – Facts or fiction?" Up Time Magazine, 8, 2000, p.113-117.
[10] P.W. Tse, Y.H Peng and R.Yam, "Wavelet analysis and envelope detection for rolling element bearing fault diagnosis – Their effectiveness and flexibilities," J Vibration and Acoustics, ASME, 123, 2001, p.303-310.
[11] W. Youshang, S. Quio and L. Xiaolei, "The application of wavelet transform and artificial neural networks in machinery fault diagnosis," Proceedings of ICSP, 1996.p. 1609-12.
[12] I.E. Alguindigue, A.L. Buczak and R.E. Uhrig, "Monitoring of rolling element bearings using artificial neural networks," IEEE Transactions on Industrial Electronics, Vol.40, No.2, April1993.p.209-217.
[13] J.D. Wu and C.H. Liu, "Investigation of engine fault diagnosis using discreet wavelet transform and neural network," Expert Systems with Applications, 2007, doi:10.1016/J.eswa.2007.08.021.
[14] S. Rajakarunakaran, P. Venkumar, D. Devaraj and K.S.P Rao, "Artificial neural network approach for fault detection in rotary system," Applied Soft Computing, 8, 2008, p.740-8.
[15] J.A. Anderson, 'A simple neural network generating an interactive memory' Mathmetical Bioscience, 14, 1972, pp. 197-220.
[16] M.T. Hagan, H.B. Demuth and M. Beale, Neural Network Design, PWS Publishing Company, Boston, 2002.
[17] D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning representations by back propagating errors," Nature, 1986, 323, p. 533-36.
[18] M.T. Hagan and M. Menhaj, "Training feed forward networks with the Marquardt algorithm," IEEE transactions on Neural Networks, 5(6), 1994.
[19] M. Zacksenhouse, S. Braun and M. Feldman, "Toward helicopter gearbox diagnostics from a small number of examples," Mechanical systems and Signal Processing, 14(4), 2000, 523-43.
[20] S.R. Gunn, "Support vector machines for classification and regression," technical report, University of Southampton, 1998.
[21] G. Goudong, Z.S. Li and K.L. Chan, "Support vector machines for face recognition," Image and Vision Computing, 19, 2001, 631-8.
[22] O. Barzilay and V.L. Brailovsky, "On domain knowledge and feature selection using a support vector machine," Pattern Recognition Letters, 20(5), 1999, 475-84.
[23] M. Ge, R. Du, G.C. Zhang and YS. Xu, "Fault diagnosis using a support vector machine with application in sheet metal stamping operations," Mech. Systems and Signal Processing, 18, 2004, 143-159.
[24] B. Samantha, "Gear fault detection using artificial neural networks and support vectors machines with genetic algorithms," Mech. Systems and Signal Processing, 18, 2004, 625-44.
[25] Q. Hu, Z. He, Z, Zang and Y. Zi, "Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble," Mech. Systems and Signal Processing, 21, 2007, 688-705.
[26] Y. Yang, D. Yu and J. Cheng, "A fault diagnosis approach for roller bearing based on envelop spectrum and SVM," Measurement, 40, 2007, 943-50.
[27] V.N. Yapnic, Statistical Learning Theory, John Willy, New Yourk, 1998.
[28] D.E. Newland, "Wavelet Analysis of vibration, Part1 :Theory," J Vib. and acoustics, 116, 1994. p. 409-24
[29] W. J. Wang and PD McFadden, "Application of wavelets to gearbox vibration signals for fault detection," J of Sound and Vib., 192(5), 1996. p. 927-39.
[30] Z.K. Peng, PW. Tse and F.L Chu, "A comparative study of improved Hilbert-Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing," Mech. System and Signal Processing, 19(2005), p. 974-988.
[31] V.J. Samar, A. Bopardikar, R. Rao and K. Swartz, "Wavelet analysis of neuroelectric waveforms: A conceptual tutorial," Brain Language, 66, 1999, 7-60.
[32] S.G. Mallat. "A theory for multiresolution signal decomposition: The wavelet representation," IEEE Trans Pattern Anal Machine Intelligence", 11(7),1989, p.674-93.

International Science Index, Mechanical and Mechatronics Engineering Vol:2, No:7, 2008 waset.org/Publication/9372