Comparison of wavelet based denoising schemes for gear condition monitoring: An Artificial Neural Network based Approach

Rounaq Ahmed¹, Srinivasa Pai P², N.S.Sriram³, Vasudeva Bhat⁴

¹Department of Mechanical Engineering, PACE, Mangalore, VTU, Karnataka, India.  
²Department of Mechanical Engineering, NMAMIT, Nitte, VTU, Karnataka, India.  
³Department of Mechanical Engineering, VVIT, Mysore, VTU, Karnataka, India.  
⁴Department of Mechanical Engineering, SIT, Mangalore, VTU, Karnataka, India.  

E-mail: rounaq77@gmail.com

Abstract. Vibration Analysis has been extensively used in recent past for gear fault diagnosis. The vibration signals extracted is usually contaminated with noise and may lead to wrong interpretation of results. The denoising of extracted vibration signals helps the fault diagnosis by giving meaningful results. Wavelet Transform (WT) increases signal to noise ratio (SNR), reduces root mean square error (RMSE) and is effective to denoise the gear vibration signals. The extracted signals have to be denoised by selecting a proper denoising scheme in order to prevent the loss of signal information along with noise. An approach has been made in this work to show the effectiveness of Principal Component Analysis (PCA) to denoise gear vibration signal. In this regard three selected wavelet based denoising schemes namely PCA, Empirical Mode Decomposition (EMD), Neighcoeff Coefficient (NC), has been compared with Adaptive Threshold (AT) an extensively used wavelet based denoising scheme for gear vibration signal. The vibration signals acquired from a customized gear test rig were denoised by above mentioned four denoising schemes. The fault identification capability as well as SNR, Kurtosis and RMSE for the four denoising schemes have been compared. Features extracted from the denoised signals have been used to train and test artificial neural network (ANN) models. The performances of the four denoising schemes have been evaluated based on the performance of the ANN models. The best denoising scheme has been identified, based on the classification accuracy results. PCA is effective in all the regards as a best denoising scheme.

1. Introduction

Gearboxes are commonly used mechanical component in industries. Gear fault results in serious damage to the gearbox as well as machine elements connected to it. The gear faults may sometimes halt the entire system [1]. Gear fault diagnosis has gained importance in industries. Vibration signal analysis has been
commonly used for reliable online condition-monitoring (CM) requirements in industries to prevent serious damages. The rotary machineries including gears exhibits non stationary phenomenon, and hence vibration signal analysis using time frequency domain and WT has been widely used [2]. During signal processing noise has to be suppressed and the clear signal is to be taken for analysis. WT is one of the powerful approaches to denoise the signal since decomposition & reconstruction of signal is done in such a way that the result contains the desired information concentrated in fewer values of wavelet coefficients [3]. PCA has been used to denoise bearing vibration signals [4], noisy image [5] and has given good results. PCA decorrelates the signal and transforms the bigger dataset to smaller one with concentrated energy and retaining original signal in the optimal principal component dataset. This optimal dimensional reduction property of PCA helps to remove noise and hence denoising the signal [6,7]. EMD has been used to denoise in a variety of applications [8, 9] and also to detect tooth faults in gearbox [10]. NC as a denoising method has been used in various applications such as image denoising, [11] fault detection in rotating machinery [12]. AT has been used to denoise gear vibration signals and fault diagnosis and has been used extensively in CM of gears [13-17].

This work is an attempt at using ANN for comparing the performance of selected three wavelet based denoising schemes (PCA, EMD and NC), with an extensively used denoising scheme (AT) for denoising gear vibration signals. These schemes have been compared using SNR, Kurtosis and RMSE. The best denoising scheme based on fault identification capability using FFT is checked. Further to ascertain the effectiveness of selected denoising scheme, statistical features has been extracted from the denoised signal and used in ANN model. The performance of ANN models in terms of classification accuracies on training and test data has been evaluated. It is found that PCA results in good classification accuracies, establishing its effectiveness for gear fault diagnosis using vibration signals.

2. Gear test rig and data acquisition

The experimental set up used for acquiring gear vibration signals is shown in figure 1. Maruti 800 gearbox with suitable modification has been used to develop gear test rig. The gear test rig mainly consists of a shaft of diameter 28.04mm which is supported between bearings. The shaft is driven by a motor of maximum speed of 1440 rpm and rated power 1.1 kW. The variation in the speed is achieved by gear shifting mechanism. The gearbox can be run with the four speed ratios 3.58, 2.16, 1.54 and 0.9 respectively. The test rig has belt-drum system as a loading unit. The vibration data has been sampled by using a KD37V accelerometer, which is mounted on the casing of the gearbox as shown in figure 1. The accelerometer has a sensitivity of 50 ± 20% mv/g and range of 120g (1170 m/s²). The vibration signals has been acquired by a DAQ system and the data has been transferred to the PC with Lab VIEW software at a sampling rate of 66 kHz and stored as text file for further analysis.

Figure 1. Experimental Setup. Figure 2. GG, HTR and FTR Gears.
Experiments have been carried out under no load and loaded conditions for speeds of 396.28 rpm (1st speed), 655.586 rpm (2nd speed), 1065.266 rpm (3rd speed) and 1577.77 rpm (4th speed). The vibration signals have been acquired in radial direction. The three gear conditions are GG, half teeth removed (HTR) and full teeth removed (FTR) were selected as defective condition as shown in figure 2.

3. Wavelet based denoising schemes

3.1. Principal component analysis

PCA reduces the large dimensional data to low dimensional data, while retaining most significant information [18]. Romero applied PCA methodology to the electrocardiogram (ECG) signals and then inverted the selected sub sets and obtained improved SNR [19]. PCA gives ‘n’ new principal components which can be selectively transformed and reverted with the selection of components filtering the noisy components. The noise in the signal can be reduced by suppressing the smaller wavelet coefficients which usually has lesser energies and considering large wavelet coefficients with high energies [20]. PCA has a capability to preserve the several most significant principal components. The algorithm for PCA based denoising scheme is given as [7]:

1. Compute the mean of the signal vector \( \mu \)
   \[
   \mu = \frac{1}{k} \sum_{i=1}^{k} x_i
   \]
   where \( k \) is number of samples, \( x \) is the signal
2. Find the covariance matrix, \( C \)
   \[
   C = \frac{1}{k} \sum_{i=1}^{k} (x_i - \mu)(x_i - \mu)^T
   \]
   where ‘T’ represents matrix transposition
3. Compute eigenvalues \( \lambda_i \) and Eigen vectors \( \nu_i \) of covariance matrix \( C \)
   \[
   (C - \lambda_i I)\nu_i = 0
   \]
   \( i = 1, 2, 3 \ldots \)
4. Estimating high-valued eigenvectors
   Arrange the Eigenvalues \( \lambda_i \) in descending order
5. Map the original signal with eigenvalues
   \[
   u = x \ast \lambda
   \]

3.2. Empirical Mode Decomposition

Empirical Mode Decomposition method decomposes the signal into a finite number of sub components known as intrinsic mode functions (IMFs). These IMFs are the representation of oscillatory mode of a particular signal obtained by systematic processing known as sifting [21]. And it should satisfy the following two properties: a) “1” should be the highest difference between the extremas and the number of zero crossings. b) At any given point, envelope mean created by the minima and maxemas should be zero [21]. For a given signal \( x(t) \), the algorithm for performing sifting is as follows:
(i) All the maximas and minimas for signal \( x(t) \) is to be identified.
(ii) Minima, ending up with a signal \( x_{min}(t) \), and likewise between maximas to get \( x_{max}(t) \) is to be interpolated.
(iii) Average between these two envelopes is to be calculated:
   \[
   x_{ave}(t) = (x_{min}(t) + x_{max}(t) - \lambda)
   \]
   \( i = 1, 2, 3 \ldots \)
(iv) Detail \( d_i(t) \) is to be extracted, which is given as input for the next sifting iteration
   \[
   d_i(t) = x(t) - x_{ave}(t)
   \]
(v) The criterion to stop the number of sifting iterations is used in order to make sure that the IMF components retain required physical sense of both amplitude and frequency modulation. This is achieved by limiting the standard deviation (SD) between two consecutive sifting iteration results. If $k$ numbers of sifting iterations are carried out, SD is represented by equation (7).

$$SD = \sum_{i=1}^{k} \left[ \frac{|d_{k+1}(t) - d_{k}(t)|}{d_{k+1}(t)} \right]^2 \tag{7}$$

Usually the value of SD is adjusted between 0.2 and 0.3.

(vi) Once $d_0(t)$ is accepted as first IMF, $h_1(t)$, the residue is calculated as:

$$r_1(t) = x(t) - d_0(t) \tag{8}$$

$$h_1(t) = d_0(t) \tag{9}$$

In order to extract second IMF, $r_1(t)$ is given as the input to the sifting process in the next round. The EMD process can be halted when the residue $r_n(t)$ reaches a monotonic function as no more IMFs can be extracted further.

(vii) As signal $X(t)$ undergoes N rounds of sifting process, N IMF sets will be decomposed, residue signal is given as:

$$x(t) = \sum_{k=1}^{N} h_k(t) + r_n(t) \tag{10}$$

Equation (10) shows that the signal decomposed by EMD process can be reconstructed using IMF components $h_k(t)$ and the residue signal $r_n(t)$.

3.3. Neighcoeff Coefficient

During conventional denoising, thresholding is done without considering the effect of neighboring coefficients. But in NC denoising, the size of neighboring coefficient has to be chosen before thresholding [22]. This method requires a pre-filter before wavelet decomposition and a post-filter after wavelet reconstruction. The signal of interest can be given as:

$$D_{j,k} = D_{j,k} + E_{j,k} \tag{11}$$

$$D_{j,k} = (d_{j,k}^{(1)}, d_{j,k}^{(2)})^T \tag{12}$$

$E_{j,k}$ has multivariate normal distribution $\mathcal{N}(0, V_j)$. The matrix $V_j$ is the covariance matrix for the error term of level $j$. By using the standard transform

$$\theta_{j,k} = D_{j,k}V_j^{-1}D_{j,k} \tag{13}$$

The matrix $V_j$ is defined as follows:

$$V_j = \begin{cases} \frac{1}{a_1} & b_1 - b_2 \\ \frac{b_1 - b_2}{(b_1 + b_2)a_1a_2} & \frac{1}{a_1a_2} \end{cases} \tag{15}$$

where

$$a_1 = \sqrt{\text{mad} (\text{row}_1)}$$

$$a_2 = \sqrt{\text{mad} (\text{row}_2)}$$

$$b_1 = \text{mad} (a_1 \text{row}_1 + a_2 \text{row}_2)$$

$$b_2 = \text{mad} (a_1 \text{row}_1 - a_2 \text{row}_2)$$

$$\text{mad} (y) = 1.4826 \text{median} (\text{abs} (y - \text{median} (y)))$$
The two rows 1, 2 represents the rows of multi-wavelet coefficients

\[ S^2_{jk} = \sum_{n=N_0}^{N} \theta^2_{j,k,n}, \quad N = N_0 - j \quad \cdots \cdots \cdots \cdots (16) \]

In the above equation 16, selection of \( N_0 \) is as per the signal duration of the feature and the wavelet support. Here \( S^2_{jk} \) is the local energy, and the size of neighbour is \((2N+1)\) which varies with the level. The decomposition and down sampling results reduced coefficients dependence with the increase in level. So considering all the parameters and equations above the thresholding formula can be given as

\[
D_{j,k} = \begin{cases} 
    D_{j,k} \left(1 - \frac{\alpha \mu^2}{S^2_{j,k}}\right) & \text{if } S^2_{j,k} \geq \alpha \mu^2 \\
    0 & \text{otherwise} \end{cases} \quad \cdots \cdots \cdots \cdots (17)
\]

where \( \mu \) is given as \( 2\log n \) and \( \alpha \) is parameter which is used to adjust the threshold value, normally \( \alpha \) value is taken as \( \alpha = \frac{(2N+1)}{3} \).

3.4. Adaptive Threshold

In wavelet decompositions we generally assume that the fault energy is concentrated among the few wavelet coefficients. But it is not so, as the energy of noise occupies almost all the wavelet coefficients. Setting of low threshold value leads to retaining of whole signal along with noisy coefficients. If high threshold value is set, then many detail coefficients which are useful may be removed. So a proper threshold value has to be calculated and adapted in order to retain the signal and remove only the noise content. [23]

4. Use of wavelet based denoising schemes

The raw vibration signals acquired from the customized gear test rig has been analyzed using a customized MATLAB programme. Discrete Wavelet Transform (DWT) and Daubechies 8 mother wavelet has been used.

![Figure 3](image1.png)  ![Figure 4](image2.png)

**Figure 3.** FFT for GG signal by four denoising schemes.  
**Figure 4.** FFT for FTR Gear signal by four denoising schemes.

Figure 3 shows the FFT plot for GG fourth speed, loaded condition, denoised by four denoised schemes. FFT plot for FTR under fourth speed, loaded condition, denoised by four denoised schemes is given by
From the figure 3 and 4 it is clear that the amplitude for FTR increases for PCA and AT denoised schemes at GMF 710 Hz and classifies the gear faults. The fault identification is difficult in case of EMD and NC, since there is not much variation in the amplitude from GG to FTR at GMF 710 Hz. Hence PCA and AT denoising schemes are helpful in gear fault detection.

SNR is the ratio of power of signal to the power of noise. SNR makes the comparison of a desired signal with the background noise level. It is often used to compare denoising schemes [24]. Increase in SNR increases performance of denoising and is given as:

$$SNR = 10 \log \frac{\sum x_i^2}{\sum (d_i - x_i)^2}$$  \hspace{1cm} (18)

where x represents raw signal, d represents denoised signal and N represents the signal length.

The shape of amplitude distribution describes the peaked and flatness of the data and hence, can be used as a data descriptor. Vibration signal containing higher valued sharp peak has distribution function sharper. Usually damaged gear box produce these types of signals. Hence kurtosis value for damaged gearbox will be higher than the good gearbox and mathematical equation is given by (19) [25].

$$K = P \sum (x_i - \bar{x})^4 \left( \sum (x_i - \bar{x})^2 \right)^{-2}$$  \hspace{1cm} (19)

where K is kurtosis, P is the number of points in the time history of signal x, x is the i\textsuperscript{th} point in the time history of signal x. Thus kurtosis for a signal can be derived by the fourth centralized moment, normalized by the variance square. Along with SNR, RMSE has also been used as a significant tool for the comparison of denoising schemes [24]. RMSE is the sample standard deviation which considers differences among predicted and observed values. As RMSE value decreases the performance of denoising increases and is shown by equation (20)[24].

$$RMSE = \left( \frac{\sum (d_i - x_i)^2}{N} \right)^{\frac{1}{2}}$$  \hspace{1cm} (20)

where x represents raw signal, d represents denoised signal and N represents the signal length.

SNR, Kurtosis and RMSE values for four denoising schemes and for all the gear conditions (GG, HTR and FTR) for fourth speed loaded condition is given in Table 1. PCA denoised signal is having highest SNR, highest kurtosis value and lowest RMSE for all conditions and values of AT denoised signal are closer to PCA denoised signal. EMD and NC denoised signals are having comparatively lesser SNR, Kurtosis values and has higher RMSE values. The same trend is followed by SNR, Kurtosis and RMSE for no load, for all the gear conditions.
Table 1. SNR, Kurtosis and RMSE values for four denoising schemes.

| Denoising Scheme | SNR  | Kurtosis  | RMSE  |
|------------------|------|-----------|-------|
|                  | GG   | HTR       | FTR   | GG   | HTR       | FTR   |
| PCA              | 75.2255 | 77.6072 | 86.7492 | 7.9891 | 13.2301 | 110.9242 | 0.00072 | 0.0068 | 0.0059 |
| EMD              | 4.0023 | 1.2798 | 1.1293 | 6.2879 | 12.0089 | 84.0883 | 0.02500 | 0.2451 | 0.3739 |
| NC               | 8.9233 | 5.6777 | 12.6174 | 6.9123 | 12.2365 | 93.1260 | 0.01980 | 0.2173 | 0.2415 |
| AT               | 74.2168 | 73.3893 | 83.8016 | 7.9710 | 13.1569 | 110.5836 | 0.00075 | 0.0072 | 0.0062 |

5. ANN based comparison of different denoising schemes

ANN are parallel distributed processing systems constructed with simple processing units called neurons, based on the concept of human brain learning process. [26,27].

5.1. Multilayer perceptron neural network

MLPNN consists of an input layer of source nodes, one or more hidden layers of computation nodes and an output layer. General MLPNN structure is shown in figure 5. Each layer has greater than one node. Node in each layer is connected to all the nodes in the neighboring layers. The connections comprises of individual weights called synaptic weights which multiplies with the node values of previous layer. The data dimensions of input and output of an ANN determine the number of nodes in the input and output layers. The number of hidden layers and their nodes can be found heuristically or by geometric pyramid rule given by.

\[ h = \sqrt{mn} \] (21)

where h , m and n are the number of nodes in hidden layer, input layer and output layer respectively.

The classification power of a MLPNN depends upon the number of hidden layers as well as nodes. Optimum number of hidden layers as well as nodes works fine and more than the optimum number leads to over fitting of the classifier and substantially increases the computation time and effort.[26,27].

The denoised signals obtained are decomposed to three levels using DWT. Eight statistical features namely standard deviation, root mean square (RMS), kurtosis, skewness, crest factor, mean, shape factor, and peak have been extracted from detail coefficients from all three levels and also from approximate coefficients at the third level. 32 features obtained from each bin forms a single pattern. For three gear conditions and four speeds with loaded and no load conditions and for 52 bins a total of 1248 patterns (3X4X2X52) were extracted. The feature set matrix consists of 32 features and 1248 patterns.

In this work, MLPNN is used to classify the three gear conditions namely GG, HTR and FTR according to the features extracted from vibration signals denoised by four schemes. Binary coding scheme has been used for implementing three conditions of gear as outputs of MLPNN, i.e. for GG (1 0 0), HTR (0 1 0) and FTR (0 0 1). Each feature is normalized, in order to get values between 0 and 1. The patterns of the matrix are thoroughly mixed, out of which 998 patterns (80%) are in the training data set and the remaining 250 patterns (20%) in the test data set.
The number of neurons in the hidden layer has been incremented in steps of 1 and maximum number of training epochs is set to 1000. Only one hidden layer is used. Different sets of learning rate (LR) and momentum rate (MR) have been studied. The sigmoid activation function has been used in hidden as well as output neurons. A mean square error (mse) of $10^{-6}$ and a minimum gradient of $10^{-10}$ are used. The training process stops if any one of the conditions is met.

The initial biases and weights of the network are randomly fixed. The MLPNN implementation is carried out using Neural Network Toolbox of MATLAB 2015. Figure 5 shows the structure of the MLPNN used, where $x_1, x_2, \ldots, x_n$ are the inputs (features), $n$ are the number of nodes in the hidden layer, and $w_{ji}$ and $w_{oj}$ are the connection weights between the input-hidden layers and the hidden-output layers, respectively.

The performance of MLPNN considering features extracted from the four denoised signals is shown in figure 6 for training and test data. PCA stands first among all denoising schemes with 98.8% and 96.84% accuracy for training and test data respectively. For 6 neurons and 1000 epochs PCA has the lowest mse value 0.0058 compared to other denoising schemes EMD, NC and AT with mse values 0.0266, 0.0186 and 0.0319 respectively.

Figure 5. Structure of the MLPNN architecture.

Figure 6. Prediction accuracies of ANN model for four denoising schemes.

6. Discussion

The focus of this work is to evaluate three wavelet based denoising schemes (PCA, EMD and NC) and commonly used denoising scheme (AT) for denoising of gear vibration signal. Initially these schemes were used and fault identification is carried out using FFT for different gear conditions. PCA and AT based denoising scheme is found to be good in identification of gear faults. Based on SNR, Kurtosis and RMSE comparison for these denoising schemes, PCA gives the best performance as shown in Table 1. In all the cases AT based denoising scheme is close to PCA based denoising scheme.
Wavelet based denoising approach is powerful as it retains the desired information by retaining fewer coefficients containing the information and removing the noise effectively. PCA performs dimensionality reduction reducing bigger dataset to smaller data set and retains the important principal components, which in turn retains the original signal. PCA and AT both identifies the fault using FFT analysis and further to select the best denoising scheme ANN has been used.

The features extracted from the denoised signals are given as inputs to ANN. The prediction accuracies of ANN model developed based on the data evolved from four different denoising schemes is compared in figure 6. Among all the schemes, features extracted from PCA based denoising scheme shows good performance compared to other denoising schemes with 98.8% accuracy on training data and 96.84% accuracy on test data. The proposed method of selecting best denoising scheme using ANN is effective. Hence PCA based denoising scheme can be effectively used to denoise the gear vibration signals compared to other denoising schemes.

7. Conclusions
This work is related to evaluation of four different wavelet based denoising schemes for denoising gear vibration signals. The four denoising schemes (PCA, EMD, NC and AT) were selected from an extensive literature survey in the area of denoising. AT based denoising scheme is a widely used denoising scheme for gear vibration signals. The AT has been compared with other three schemes. PCA and AT based denoised signals can identify gear faults effectively. Based on SNR, Kurtosis and RMSE, PCA denoising scheme gives good result.

The performance of these schemes has been evaluated, considering the features extracted from the denoised signals and using it in an ANN model. Though AT gives comparable performance with PCA with regard to the accuracy needed for the training and test data, PCA gives the best results among all denoising schemes with 98.8% and 96.84% accuracy for training and test data respectively. Thus PCA can be selected as best denoising scheme for gear vibration signals and ANN can be effectively used to evaluate the different denoising schemes.

References
1. Sung, C. K., H. M. Tai, and C. W. Chen. "Locating defects of a gear system by the technique of wavelet transform." Mechanism and machine theory 35.8 (2000): 1169-1182.
2. N. Saravanan, K.I. Ramachandran, Fault diagnosis of spur bevel gear box using discrete wavelet features and Decision Tree classification, Expert Syst. Appl. 36 (2009) 9564-9573.
3. Nieslony, Arthur. "Wavelet-Based Methods for Denoising." Internship Report (2005).
4. Xiang, Jiawei, Yongteng Zhong, and Haifeng Gao. "Rolling element bearing fault detection using PPCA and spectral kurtosis." Measurement 75 (2015): 180-191.
5. Manjón, José V., Pierrick Coupé, and Antonio Buades. "MRI noise estimation and denoising using non-local PCA." Medical image analysis 22.1 (2015): 35-47.
6. Zhang, Lei, et al. "PCA-based spatially adaptive denoising of CFA images for single-sensor digital cameras." Image Processing, IEEE Transactions on 18.4 (2009): 797-812.
7. Zhang, Lei, et al. "Two-stage image denoising by principal component analysis with local pixel grouping." Pattern Recognition 43.4 (2010): 1531-1549.
8. Tian, Pengfei, et al. "The Application of EMD-CIIT Lidar Signal Denoising Method in Aerosol Detection." Procedia Engineering 102 (2015): 1233-1237.
9. Guo, Chencheng, et al. "Adaptive noise cancellation based on EMD in water-supply pipeline leak detection." Measurement (2015).
10. Jin, Shuai, Jong-Sik Kim, and Sang-Kwon Lee. "Sensitive method for detecting tooth faults in gearboxes based on wavelet denoising and empirical mode decomposition." Journal of Mechanical Science and Technology 29.8 (2015): 3165-3173.

11. Om, Hari, and Mantosh Biswas. "A generalized image denoising method using neighbouring wavelet coefficients." Signal, Image and Video Processing 9.1 (2015): 191-200.

12. Wang, Yanxue, Zhengjia He, and Yanyang Zi. "Enhancement of signal denoising and multiple fault signatures detecting in rotating machinery using dual-tree complex wavelet transform." Mechanical Systems and Signal Processing 24.1 (2010): 119-137.

13. TANG, Jin-yuan, et al. "Wavelet-based vibration signal denoising with a new adaptive thresholding function [J]." Journal of Vibration and Shock 7 (2009): 027.

14. Wenbin, Li, et al. "Application of Multiwavelet Adaptive Threshold Denoising in the Fault Diagnosis of Gearbox." Measuring Technology and Mechatronics Automation (ICMTMA), 2011 Third International Conference on. Vol. 1. IEEE, 2011.

15. Belsak, Ales, and Joze Flasker. "Adaptive wavelet transform method to identify cracks in gears." EURASIP Journal on Advances in Signal Processing 2010 (2010):6.

16. Jiang, Yonghua, et al. "Feature extraction method of wind turbine based on adaptive Morlet wavelet and SVD." Renewable energy 36.8 (2011): 2146-2153.

17. Liu, Xiangui, and Wuwei Chen. "Approach to Denoising of Local Adaptive Algorithm Based on DT-CWT and Its Application to Auto Transmission Fault Diagnosis." Fuzzy Systems and Knowledge Discovery, 2008. FSKD'08. Fifth International Conference on. Vol. 4. IEEE, 2008.

18. Anon, Principal Component Analysis (PCA) Basics.BRUKER www.bdal.com

19. Romero, Inaki. "PCA-based noise reduction in ambulatory ECGs." 2010 Computing in Cardiology. IEEE, 2010.

20. Romero, I. "PCA and ICA applied to Noise Reduction in Multi-lead ECG."Computing in Cardiology, 2011. IEEE, 2011.

21. Chacko, Ani, and Samit Ari. "Denoising of ECG signals using Empirical Mode Decomposition based technique." Advances in Engineering, Science and Management (ICAESM), 2012 International Conference on. IEEE, 2012.

22. Wang, Xiaodong, Yanyang Zi, and Zhengjia He. "Multiwavelet denoising with improved neighboring coefficients for application on rolling bearing fault diagnosis." Mechanical Systems and Signal Processing 25.1 (2011): 285-304.

23. Wenbin, Li, et al. "Application of Multiwavelet Adaptive Threshold Denoising in the Fault Diagnosis of Gearbox." Measuring Technology and Mechatronics Automation (ICMTMA), 2011 Third International Conference on. Vol. 1. IEEE, 2011.

24. Vijay, S. G., et al. "Evaluation of effectiveness of wavelet based denoising schemes using ANN and SVM for bearing condition classification."Computational intelligence and neuroscience 2012 (2012):16.

25. P.Vecer,M.Kreidl,R.Smid,"Condition Indicators for Gearbox Condition Monitoring Systems," Acta Polytechnica Vol.45 No 6/2005.36-37.

26. Saravanan, N., VNS Kumar Siddabattuni, and K. I. Ramachandran. "Fault diagnosis of spur bevel gear box using artificial neural network (ANN), and proximal support vector machine (PSVM)." Applied soft computing 10.1 (2010): 344-360.

27. Rafiee, J., et al. "Intelligent condition monitoring of a gearbox using artificial neural network." Mechanical systems and signal processing 21.4 (2007): 1746-1754.