Envelope-Wavelet Packet Transform for Machine Condition Monitoring

M. F. Yaqub, I. Gondal, and J. Kamruzzaman

***Abstract*—**Wavelet transform has been extensively used in machine fault diagnosis and prognosis owing to its strength to deal with non-stationary signals. The existing Wavelet transform based schemes for fault diagnosis employ wavelet decomposition of the entire vibration frequency which not only involve huge computational overhead in extracting the features but also increases the dimensionality of the feature vector. This increase in the dimensionality has the tendency to 'over-fit' the training data and could mislead the fault diagnostic model. In this paper a novel technique, envelope wavelet packet transform (EWPT) is proposed in which features are extracted based on wavelet packet transform of the filtered envelope signal rather than the overall vibration signal. It not only reduces the computational overhead in terms of reduced number of wavelet decomposition levels and features but also improves the fault detection accuracy. Analytical expressions are provided for the optimal frequency resolution and decomposition level selection in EWPT. Experimental results with both actual and simulated machine fault data demonstrate significant gain in fault detection ability by EWPT at reduced complexity compared to existing techniques.

***Keywords*—**Envelope Detection, Wavelet Transform, Bearing Faults, Machine Health Monitoring.

I. Introduction

MACHINE condition monitoring (MCM) is crucial in all industrial processes to achieve high reliability, reduced man power and scheduled maintenance. Rotary machinery is the crucial component in almost all the industrial processes. In case of rotary machinery, malfunctioning in the operation of the bearing is the most common fault. It has been investigated that 40% of the total machine faults are because of bearing faults [1]. In the bearing faults, certain vibration patterns are generated whenever the rolling element passes the defect position. These vibration patterns vary according to the variations in the machine dynamics. The vibration characteristics frequencies generated by inner race, ball and outer race faults are given in (1)-(3) which are dependent upon the speed of rotation and geometry of the bearing [2].

$$f_{ID} = \frac{n}{2} f_{rm}\left(1 + \frac{d_{ball}}{d_{pitch}} cos\phi\right), \quad (1)$$

$$f_{OD} = \frac{n}{2} f_{rm}\left(1 - \frac{d_{ball}}{d_{pitch}} cos\phi\right), \quad (2)$$

$$f_{BD} = \frac{d_{pich}}{2d_{ball}} f_{rm}\left(1 - \left(\frac{d_{ball}}{d_{pitch}}\right)^2 cos^2\phi\right), \quad (3)$$

where f_{rm}, d_{pitch}, d_{ball}, n and ϕ represent the rotational speed, pitch diameter, ball diameter, number of balls and the contact angle as highlighted in Fig. 1.

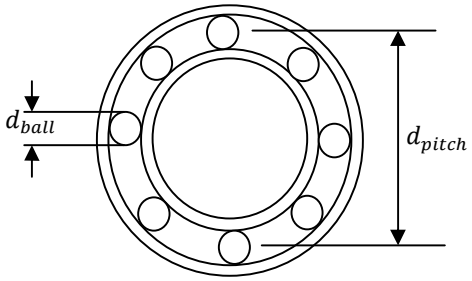

Fig. 1 Rolling element ball bearing geometry

In MCM, there are many techniques which deal with diagnosis and prognosis of machine faults [3-5] and some are also proposed by the authors [6-10]. Wavelet packet transform has been extensively used in MCM to extract the frequency domain information in case of non-stationary signals [4, 11-17]. Wavelet transform divides the overall vibration spectrum into multiple sub-bands. The entire frequency band is not enriched in fault diagnostic signatures, rather only the resonant frequency band contains dominant fault signatures as signal to noise ratio (SNR) of the vibration data is maximal at resonant frequency band [18]. The existing schemes dealing with wavelet packet transform (WPT) decompose the vibration signal to very high decomposition levels to extract the fault related information into different frequency sub-bands. Higher decomposition level results in larger number of features [4, 14]. Only a few of the features correspond to the resonant frequency band and contributes dominantly in fault diagnosis. The remaining redundant features may mislead the fault diagnostic model and decrease its accuracy in detecting fault. Moreover, large number of features cause extra computational overhead in terms of computing feature vector as well as increase the complexity of the fault diagnostic model.

In this paper, a new technique, envelope wavelet packet transform (EWPT) is proposed. In EWPT, envelope [19] of the vibration data is determined and the features are only extracted for the part of the envelope spectrum which contains fault diagnostic information. As a result reduced number of features is extracted, which achieves enhancement in the fault

M. F. Yaqub is with Monash University, Australia, on leave from Department of Electrical Engineering, University of Engineering and Technology, Lahore. (Phone: +61-3-512-26568; e-mail: farrukh.yaqub@ monash .edu).

I. Gondal is with Gippsland School of Information Technology, Monash University, Australia. (e-mail: iqbal.gondal@monash.edu).

J. Kamruzzaman is with Gippsland School of Information Technology, Monash University, Australia. (e-mail: joarder.kamruzzaman@monash. edu).

detection accuracy and reduced complexity. Moreover, analytical expressions are provided for the optimal selection of the decomposition level, which provides requisite classifiable information to extract the features for the fault diagnostic model.

The paper is organized as follows, Section II presents framework for EWPT, Section III provides results and substantiates the need for EWPT in comparison to the techniques based on WPT. Section IV provides concluding remarks for the proposed scheme.

II. Framework for Envelope-Wavelet Packet Transform

This section provides framework for the fault diagnostic model based on EWPT. EWPT follows a hierarchical paradigm as illustrated in Fig. 2. First, an envelope of the digitized vibration data is determined. There are different techniques for envelope detection, in this paper, Hilbert transform is used [20]. The envelope of signal is extracted such that it contains only the band which corresponds to the fault signatures. Second, Wavelet packet transform is applied on the vibration data corresponding to the envelope signal. The decomposition level of the wavelet is selected such that the frequency bands representative of the different faults lie in different wavelet decomposition nodes. It helps the fault diagnostic model to correctly classify the type of fault. Third, a diagnostic model is built using the extracted data.

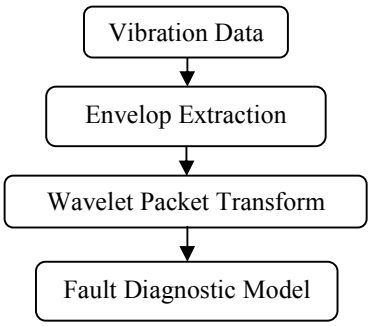

Fig. 2 Envelope-Wavelet Packet Transform (EWPT)

A. Envelope Detection:

Envelope detection is extensively used technique to determine bearing faults [19]. EWPT validates that the envelope detection together with wavelet packet transform enhances the robustness of the fault diagnostic model. In EWPT, the envelop of the signal is determined based on Hilbert's transform [20] as given in (4):

$$f(t)_{env}{}^2 = f(t)^2 + \mathcal{H}\{f(t)\}^2, \quad (4)$$

where $\mathcal{H}\{f(t)\}$ represents the 'Hilbert' transform of the signal $f(t)$.

B. Wavelet Packet Transform

In order to extract the time-frequency information in the non-stationary vibration signal, digitized vibration data are decomposed using WPT [4, 11-17]. WPT helps in investigating the frequency contents of the vibration data in different frequency ranges, i.e., nodes. Among different time-frequency domain signal processing techniques, e.g., discrete Fourier transform (DFT), short time Fourier transform (STFT) and wavelet packet transform (WPT), WPT can be used for comprehensive analysis of non-stationary vibration signal to reliably extract its time and frequency domain contents [11]. DFT of a non-stationary signal $x[n]$ (5) does not exploit the variation in frequency contents with respect to time. Rather it averages out the frequency content over the whole signal range [11].

$$X(k) = \sum_{n=0}^{N-1} x[n]\, e^{-j\left(\frac{2\pi}{N}\right)kn},\ k = 0,1,\dots(N-1). \quad (5)$$

The shortcomings of DFT can be overcome by STFT (6), but STFT suffers from the problem that it yields the same time and frequency resolution for both low and high frequencies. The time and frequency resolution remains same because window size $\omega[n]$ remains constant throughout the analysis [11].

$$X(m,k) = \sum_{n=0}^{N-1} x[n]\omega[m-n]e^{-j\left(\frac{2\pi}{N}\right)kn},$$
$$k = 0,1,\dots,(N-1). \quad (6)$$

In order to overcome the drawback of fixed time-frequency resolution in STFT, WPT can be used which has the tendency to perform multi-resolution analysis. Digitized vibration data are passed through high pass $h[n]$ and low pass $g[n]$ Quadrature Mirror Filters (QMFs) (7)-(8) and then down sampled. QMFs are finite impulse response (FIR) filters or infinite impulse response (IIR) filters [21]. Filter selection is a very crucial part in case of analysis using WT. In the proposed scheme, Daubechies (Db5) filter [14] is used which is an FIR filter. The total number of nodes at any $j-th$ decomposition level is given by (9):

$$y_{approx}[n] = x[n] * g[n],$$
$$or \qquad y_{approx}[n] = \sum_{k=-\infty}^{k=\infty} x[k] \times g[n-k]. \quad (7)$$

$$y_{detailed}[n] = x[n] * h[n],$$
$$or \qquad y_{detailed}[n] = \sum_{k=-\infty}^{k=\infty} x[k] \times h[n-k]. \quad (8)$$

$$N_j = 2^{j_{decomp}}. \quad (9)$$

C. Envelope-Wavelet Packet Transform (EWPT):

In EWPT, the envelope of the incoming vibration data is determined as in (4). In the bearing faults, the defect frequency spectrum spreads over a wide band [22, 23] due to background noise. In order to capture the dominant signatures, WPT is used along envelope detection. In EWPT, instead of decomposing the entire frequency band using WPT, the portion of the envelop signal which contains dominant fault signatures is filtered first. Filtered-envelope signal is decomposed using WPT in order to extract features for the fault diagnostic model.

EWPT is designed to diagnose three types of bearing faults i.e., inner race, outer race and rolling element with defect frequencies as (1)-(3). Figure 3 presents the generic envelope spectrum of three fault signals with defect frequencies as f_A, f_B and f_C. The spreading of the defect frequency spectrum is incorporated with bandwidth as B_F. In Fig. 3, B_{env} is the bandwidth of the envelope signal and B_G is '*guard*' band between the dominant components of the frequency spectrum with different fault frequencies. In EWPT, instead of using the entire vibration spectrum of the envelope signal for wavelet transformation, the envelope signal is passed through band pass filter with cut-off frequencies as the effective bandwidth of the envelope signal, i.e., B_{env}. It results in the reduced decomposition level of the wavelet packet transform and the number of features (sub-bands) is reduced accordingly as in (9).

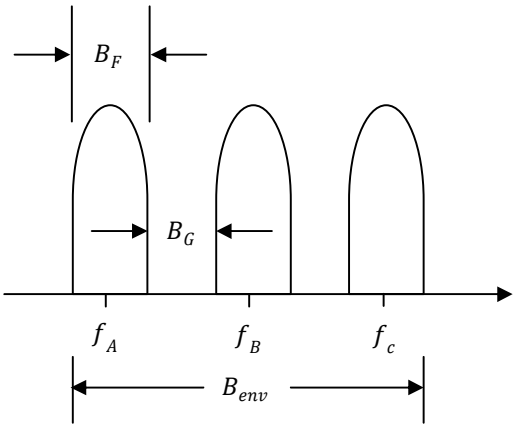

Fig. 3 Filtered-envelope of the vibration signal

The bandwidth of each of the decomposition node at the j-th decomposition level, i.e., $B_j$is given as (10):

$$B_j = \frac{B_{env}}{2^{j}decomp}. \quad (10)$$

From Fig. 3, the bandwidth of the vibration signal from the envelope detection is B_{env} prior to wavelet decomposition. In order to ensure that the information from different types of faults lies in different decomposition nodes, the frequency resolution at the j-th decomposition level, i.e., B_j should be less or equal to B_G, i.e., $B_j \leq B_G$. This limit validates that the defect frequency information of different faults would lie in different decomposition nodes.

Figure 4 elaborates the working principle of EWPT using simulated vibration data with three different defect frequencies at $f_A = 500$, $f_B = 533$ and $f_C = 566$Hz. Figure 4a plots the spectrum of the signal of the incoming vibration signal with resonant frequency band as 1500-4500Hz (approximately). The positioning of the resonant frequency band depends upon machine dynamics (transfer function) and it varies with the operating conditions. The overall bandwidth of the vibration frequency signal is half of the sampling rate, i.e., 6000Hz. The techniques based on WPT decompose the overall vibration spectrum into multiple frequency bands. It involves a large number of redundant features which do not contain dominant fault signatures for fault diagnosis. These redundant features have the tendency to mislead the fault diagnostic model due to '*curse of dimensionality*' [24] and deteriorate fault detection accuracy.

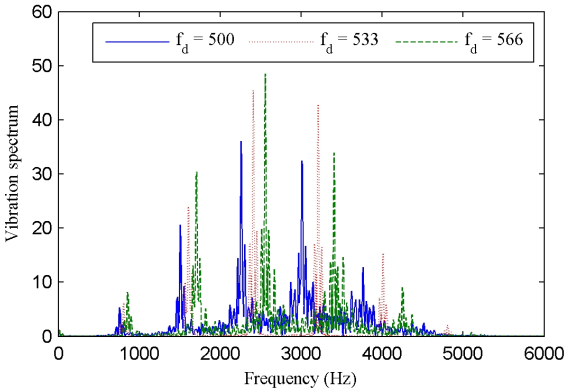

Fig. 4a Frequency spectrum of the overall vibration signal

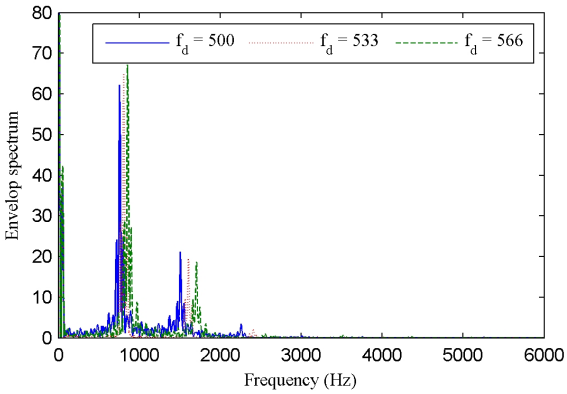

Fig. 4b Envelope spectrum of the vibration signal

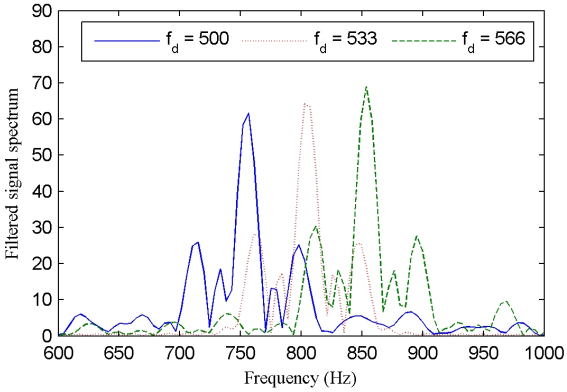

Fig. 4c Filtered envelope spectrum of the vibration signal

In EWPT, prior to extracting features based on WPT, envelope detection is applied as in Fig. 4b. Once the envelope of the signal is determined, the defect frequency region of the overall vibration is filtered out as in Fig. 4c. According to Figs. 3-4c, the WPT of the filtered-envelope of the vibration spectrum can decompose the signal such that the dominant

signatures representative of different faults lie in different frequency sub-bands. In EWPT, the WPT along with envelope detection rather than DFT is used because WPT has the tendency to provide multi-resolution analysis.

After wavelet decomposition of the filtered-envelope signal, the RMS value of wavelet decomposition nodes is computed as in (11). RMS value based feature extraction is extensively used as feature extraction metrics in the literature [4, 12, 13]. In (11), $x_{k,i,j}$ represents the $k-th$ coefficient of $i-th$ decomposition node at $j-th$ decomposition level.

$$X_{i,j} = \sqrt[2]{\frac{1}{N}\sum_{k=1}^{N} x_{k,i,j}^2}\,. \tag{11}$$

D.Fault Diagnostic Model:

Support vector machine is extensively used technique in machine fault diagnosis and prognosis [25]. It maps the data into higher dimensional space and resolves the problem of decision boundary in low dimensional space. Figure 5 gives the orientation of the optimal hyper-plane in binary classification problem. The data points nearest to the decision boundary are called, support vectors, and help in deciding the decision boundary of the hyper plane, represented as sv_1, sv_2, sv_3, and sv_4 in Fig. 5.

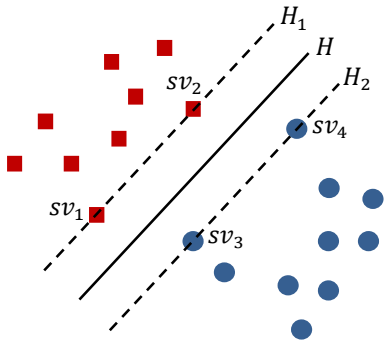

Fig. 5 Optimal Hyper-plane in SVM

Consider the training dataset $D = \{(\boldsymbol{x_i}, y_i)\}$ with an input $\boldsymbol{x_i} \in R^{Nf}$ and the corresponding output $y_i \in \{-1,1\}$. To find an optimal separating hyperplane, each input $\boldsymbol{x_i}$ is mapped to higher dimensional space via a non-linear mapping, such that $\boldsymbol{z_i} = \varphi(\boldsymbol{x_i})$. There exists a vector $\boldsymbol{w}$ and scalar b that define the separating hyperplane as $\boldsymbol{w}.\boldsymbol{z_i} + b = 0$ as in (12):

$$y_i(\boldsymbol{w}.\boldsymbol{z_i} + b) \geq 1 - \xi_i \quad where \quad \boldsymbol{w} \in R^n, b \in R, \tag{12}$$

where $\xi_i \geq 0$ are slack variables and only misclassified training samples generate non-zero ξ_i. As in Fig. 5, the optimal hyper-plane with maximal margin is equivalent to minimizing the value of $\|\boldsymbol{w}\|$ which may be defined as quadratic optimization problem as in (13):

$$min\frac{1}{2}\langle \boldsymbol{w}.\boldsymbol{w}\rangle + C\sum_{i=1}^{L}\xi_i, \tag{13}$$

where C is a constant parameter, called regularization parameter, which determines the tradeoff between maximum margin and minimum classification error. The optimization problem in (13) is quadratic programming problem and can be formulated in terms of Lagrangian multipliers [26] as in (14):

$$max\, L(\boldsymbol{\alpha}) = \sum_{i=1}^{m}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{m}\alpha_i\alpha_j y_i y_j K\langle \boldsymbol{x_i}.\boldsymbol{x_j}\rangle,$$
$$subject\ to\ \sum_{i=1}^{L}\alpha_i y_i = 0\ and\ 0 \leq \alpha_i \leq C, \tag{14}$$

where α_i is non-negative Lagrangian multipliers and K is the Kernel function which is equivalent to transforming the input feature vector $\boldsymbol{x_i}$ to higher dimensional feature space $\boldsymbol{z_i}$ as in (15):

$$K(\boldsymbol{x_i}, \boldsymbol{x_j}) = \langle \varphi(\boldsymbol{x_i}), \varphi(\boldsymbol{x_j})\rangle. \tag{15}$$

There are many choices of Kernal's such as RBF (16), Polynomial (17) and Hyperbolic tangent (18):

$$K(\boldsymbol{x_i}, \boldsymbol{x_j}) = e^{-\gamma\|x_i - x_j\|^2}. \tag{16}$$

$$K(\boldsymbol{x_i}, \boldsymbol{x_j}) = (\boldsymbol{x_i}.\boldsymbol{x_j} + 1)^d. \tag{17}$$

$$K(\boldsymbol{x_i}, \boldsymbol{x_j}) = tanh(\beta \boldsymbol{x_i}.\boldsymbol{x_j} + b). \tag{18}$$

In this paper, the performance of the proposed training set and feature selection scheme is measured using linear and RBF kernel which are most commonly used in the literature [3, 27].

The proposed model uses two parameters which are to be optimized, i.e., regularization parameter (C) and RBF kernel parameter (γ). In case of linear kernel, C is to be optimized and in case of RBF kernel, both C and γ are to be optimized. For parameter optimization, vibration data are divided into two subsets, training dataset (67%) and test dataset (33%). The training dataset is used for optimizing the parameters using 10-fold cross validation [27] and test dataset is used to evaluate the performance of the proposed model with the optimized parameters. In order to search for the optimal values of C and γ, C is varied from 2^{-5} to 2^{15} and γ is varied from 2^{-15} to 2^{10} [27]. The quantification metric for parameters optimization is based on maximal fault detection accuracy. The classification (i.e., fault detection) accuracy is measured in terms of correctly classified events ($E_{Accurate}$) to the total number of events (E_{Total}) as in (19):

$$Accuracy = \frac{E_{Accurate}}{E_{Total}} \times 100\ \%. \tag{19}$$

III. Experimental Results

A. Data Acquisition:

The performance of the proposed scheme is evaluated for both actual as well as simulated vibration data. This subsection presents the details on data acquisition.

1) Actual Vibration Data:

In this study, publically available vibration dataset has been used, in which different types of faults were created using electro-discharge-machining (EDM) [28]. Faulty bearings were supporting the shaft of a motor and the load was 2 HP at

the speed of 1750 rpm. The data were collected through accelerometers using 16 channels digital audio tape (DAT) recorder and sampled at the rate of 12,000 samples per second as in Fig. 6. Data for each of the inner race, ball and outer race faults were captured with different severity levels, i.e., a hole with diameter of 0.007 inch, 0.014 inch and 0.021 inch.

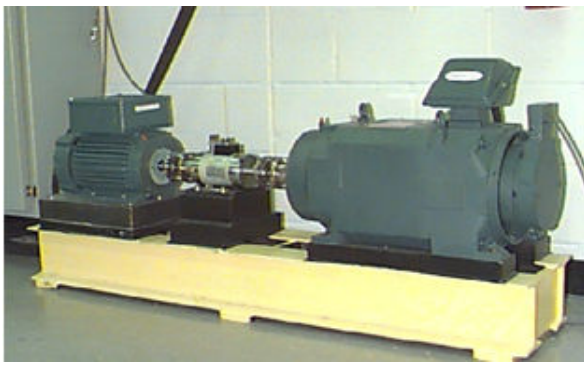

Fig. 6 Experimental setup

2) Simulated Vibration Datasets:

For simulation studies of EWPT, the vibration data is obtained from machine transfer function. The transfer function is determined by continuous hammering the rotary machine [28]. The transfer function is given by (20) and the most common input forces for the vibration signal, i.e., mass unbalance, misalignment of the shaft and bearing defect are given by (21) – (23) respectively:

$$H(s) = \frac{2.283s^4+642.6s^3+4.457\times10^8 s^2}{s^4+619.3s^3+4.42\times10^8 s^2+4.234\times10^{10}s+5.697\times10^{15}}, \quad (20)$$

$$F_{mass-unbalance} = A_1 \cos(2\pi f_s t + \theta_1), \quad (21)$$

$$F_{misalignment} = f_s \sum_{k=1}^{N_m} A_k^m cos(2\pi k f_s t + \theta_m^k), \quad (22)$$

$$F_{bearing} = f_s \sum_{k=1}^{N_b} A_k^b cos(2\pi k f_b t + \theta_b^k). \quad (23)$$

The harmonic amplitudes A, number of harmonics N, phase relationships θ, speed of rotation f_s and bearing defect frequencies f_b are adjusted for these forces to simulate the vibration data for inner race, ball and outer race faults with different severity [28].

B. Performance Evaluation of EWPT:

In order to evaluate the performance of EWPT, experiments are conducted with the test data after parameters optimization (technique presented in Subsection II-D). The dominant component of envelope bandwidth is taken as $B_{env} = 60\ to\ 200Hz$, determined from the experimental vibration datasets. The defect frequencies are determined using (1)-(3) such that $f_{ID} = 145Hz$, $f_{OD} = 90Hz$ and $f_{BD} = 117Hz$. With the given values of envelop bandwidth and defect frequencies, the optimal bandwidth of each of the decomposition node at the $j\text{-}th$ decomposition level, i.e., $B_j = 20Hz$ or optimal decomposition level is $j = 3$ using (10) and results into '8' features according to (9). The performance of the extracted features is measured using Linear and RBF kernel. Comparative performance evaluation is provided with the existing techniques based on WPT.

1) Performance in case of linear kernel:

Tables I-II list the fault detection accuracy for the linear kernel with the actual and simulated vibration datasets. Table I shows that in case of actual vibration data, the classification accuracy for the proposed EWPT reaches the maximum value at decomposition level '3' as compared to WPT in which classification accuracy is maximum at decomposition level '6'. Decomposition level '6' contains '64' features and decomposition level '3' contains '8' features. In case of WPT, a large number of features contain redundant information and misleads the classifier. The classification accuracy of EWPT is significantly higher compared to WPT in case of actual vibration data, i.e., improved by 11%. In both EWPT and WPT, detection accuracy starts decreasing beyond the decomposition level at which maximum accuracy is reached. The approaches in [4, 11-17] are based on WPT and do not contain any criterion for the optimal selection of the decomposition level. These analyses show that if decomposition level is higher than the optimal threshold, it may decrease the classification accuracy. EWPT develops a criterion for the optimal decomposition level (refer to where it is mentioned in the paper) and classification accuracy is maximized at that decomposition level, i.e., '3'.

Table II provides the classification accuracy of the fault diagnostic model in case of linear kernel with the simulated data. It shows that the classification accuracy of WPT is maximized at decomposition level '7' with '128' features whereas in case of EWPT, the classification accuracy is maximized at decomposition level '3' with only '8' features. The results with linear kernel validate that the proposed scheme enhances the classification accuracy, reduces the number of features and complexity.

TABLE I
FAULT DETECTION ACCURACY
ACTUAL VIBRATION DATA WITH LINEAR KERNEL

Decomposition Level	WPT	EWPT
1	65.5556	87.6852
2	71.6667	89.1567
3	76.6667	90.5556
4	78.3333	90.3124
5	78.8889	90.0000
6	79.5549	87.7778
7	77.5417	75.2778

TABLE II
FAULT DETECTION ACCURACY
SIMULATED VIBRATION DATA WITH LINEAR KERNEL

Decomposition Level	WPT	EWPT
1	71.5421	71.1564
2	71.5485	98.8755
3	86.8684	100.000
4	81.7451	99.5448
5	88.1254	98.9425
6	92.5965	100.000
7	100.000	100.000

2) Performance in case of RBF kernel:

Tables III list the fault detection classification accuracy using RBF kernel for the actual vibration datasets. In case of EWPT, the classification accuracy is maximized at decomposition level '3' and in case of WPT, the classification accuracy is maximized at decomposition level '5'. The fault detection accuracy for EWPT is consistently higher as compared to WPT, i.e., improved by 12% in case of RBF kernel with actual vibration data. Moreover, classification accuracy starts decreasing both in case of WPT and EWPT if decomposition level is increased beyond '5' and '3' respectively. EWPT develops intuitive and analytical expressions for optimal selection of decomposition level, i.e., $j = 3$. It validates that the technique based on WPT also require optimal selection of decomposition level otherwise classification accuracy can be deteriorated if decomposition level is lower or higher than a particular optimal value.

Table IV lists the classification accuracy in case of EWPT and WPT for the simulated vibration datasets. It shows that the classification accuracy of EWPT is maximized at decomposition level '3' as compared to WPT in which classification accuracy is maximized at decomposition level '7'. Classification results with RBF kernel are slightly improved as compared to 'linear' kernel because RBF kernel maps the data into higher dimensional space using the kernel function in (16). Overall performance of EWPT validates that it does not only enhance the fault detection accuracy but also reduces the number of features and complexity of the fault diagnostic model.

TABLE III
FAULT DETECTION ACCURACY
ACTUAL VIBRATION DATA WITH RBF KERNEL

Decomposition Level	WPT	EWPT
1	66.6667	88.5982
2	74.7222	90.5556
3	78.8889	93.3251
4	79.3512	92.5486
5	81.2156	91.6667
6	76.1111	88.0556
7	77.2222	78.3333

TABLE IV
FAULT DETECTION ACCURACY
SIMULATED VIBRATION DATA WITH RBF KERNEL

Decomposition Level	WPT	EWPT
1	70.0000	86.6667
2	81.6667	99.4444
3	87.2222	100.000
4	81.1111	99.8645
5	91.6667	99.5485
6	92.2222	100.000
7	100.000	100.000

IV. CONCLUSION

This paper proposes a novel technique, namely, envelope wavelet packet transform (EWPT). EWPT avoids the 'curse of dimensionality' by extracting features from the filtered-envelope of the vibration data on contrary to the existing schemes based on WPT, which extracts features from the overall vibration spectrum. Moreover, a criterion is devised for optimal frequency resolution which enhances the strength of fault signatures in terms of their diagnostic tendency. The results obtained both for the actual as well as the simulated vibration datasets validate that EWPT does not only reduce the dimensionality of the feature vector and complexity of the fault diagnostic model but also promises significant performance enhancement in terms of fault detection accuracy. Together the theoretical analysis and experimental results validate that EWPT has better fault diagnostic capability which justifies its massive deployment in future fault diagnostic systems.

REFERENCES

[1] J. Morel, "Vibratory monitoring and predictive maintenance," *Techniques de l'Ingénieur, Measurement and Control,* vol. RD, 2002.
[2] M. El Hachemi Benbouzid, "A review of induction motors signature analysis as a medium for faults detection," *IEEE Trans. on Ind. Electron.,* vol. 47, pp. 984-993, 2000.
[3] Q. Hu, Z. He, Z. Zhang, and Y. Zi, "Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble," *Mechanical Systems and Signal Processing,* vol. 21, pp. 688-705, 2007.
[4] K. Teotrakool, M. J. Devaney, and L. Eren, "Adjustable-Speed Drive Bearing-Fault Detection Via Wavelet Packet Decomposition," *IEEE Trans. on Instrum. and Meas.,* vol. 58, pp. 2747-2754, 2009.
[5] J. R. Stack, T. G. Habetler, and R. G. Harley, "Fault-signature modeling and detection of inner-race bearing faults," *IEEE Trans. on Industry Applications.,* vol. 42, pp. 61-68, 2006.
[6] M. F. Yaqub, I. Gondal, and J. Kamruzzaman, "Machine Fault Severity Estimation Based on Adaptive Wavelet Nodes Selection and SVM (Accepted for publication)," in *IEEE International Conference on Mechatronics and Automation*, China, 2011.
[7] M. F. Yaqub, I. Gondal, and J. Kamruzzaman, "Severity Invariant Machine Fault Diagnosis (Accepted for publication)," in *IEEE International Conference on Industrial Electronics and Application*, China, 2011.
[8] M. F. Yaqub, I. Gondal, and J. Kamruzzaman, "Resonant Frequency Band Estimation using Adaptive Wavelet Decomposition Level Selection (Accepted for publication)," in *IEEE International Conference on Mechatronics and Automation*, China, 2011.
[9] M. F. Yaqub, I. Gondal, and J. Kamruzzaman, "Severity Invariant Feature Selection for Machine Health Monitoring," *International Review of Electrical Egnineering,* vol. 6, pp. 238-248, 2011.
[10] M. F. Yaqub, I. Gondal, and J. Kamruzzaman, "Machine Health Monitoring Based on Stationary Wavelet Transform and 4th Order Cumulants (Accepted for publication)," *Australian Journal of Electrical & Electronics Engineering,* 2011.
[11] Z. K. Peng and F. L. Chu, "Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography," *Mechanical Systems and Signal Processing,* vol. 18, pp. 199-221, 2004.
[12] L. Eren and M. J. Devaney, "Bearing damage detection via wavelet packet decomposition of the stator current," *IEEE Trans. Instrum. and Meas.,* vol. 53, pp. 431-436, 2004.
[13] E. C. C. Lau and H. W. Ngan, "Detection of Motor Bearing Outer Raceway Defect by Wavelet Packet Transformed Motor Current Signature Analysis," *IEEE Trans. Instrum. and Meas.,* vol. 59, pp. 2683-2690, 2010.
[14] G. Y. Yen and L. Kuo-Chung, "Wavelet packet feature extraction for vibration monitoring," *Proceedings of IEEE International Conference on Control Applications,* 1999, pp. 1573-1578 vol. 2.
[15] F. Zhao, J. Chen, and W. Xu, "Condition prediction based on wavelet packet transform and least squares support vector machine methods," *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering,* vol. 223, pp. 71-79, 2009.
[16] L. Eren, Y. Cekic, and M. J. Devaney, "Enhanced feature selection from wavelet packet coefficients in fault diagnosis of induction motors with

artificial neural networks," in *IEEE Instrumentation and Measurement Technology Conference (I2MTC),* 2010, pp. 960-963.
[17] Z. Jianhua, Y. Zhixin, and S. F. Wong, "Machine condition monitoring and fault diagnosis based on support vector machine," in *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2010, pp. 2228-2233.
[18] I. S. Bozchalooi and M. Liang, "A joint resonance frequency estimation and in-band noise reduction method for enhancing the detectability of bearing fault signals," *Mechanical Systems and Signal Processing,* vol. 22, pp. 915-933, 2008.
[19] R. B. Randall, J. Antoni, and S. Chobsaard, "The relationship between spectral correlation and envelope analysis in the diagnossi of bearing faults and other cyclostationary machine signals," *Mechanical Systems and Signal Processing,* vol. 15, pp. 945-962, 2001.
[20] P. W. Tse, Y. H. Peng, and R. Yam, "Wavelet Analysis and Envelope Detection For Rolling Element Bearing Fault Diagnosis---Their Effectiveness and Flexibilities," *Journal of Vibration and Acoustics,* vol. 123, pp. 303-310, 2001.
[21] J. S. Walker, *A Primer on Wavelets and their Scientific Applications*. New York: Chapman &Hall/CRC, 1999.
[22] W. Changting and R. X. Gao, "Wavelet transform with spectral post-processing for enhanced feature extraction," in *Proceedings of the 19th IEEE Instrumentation and Measurement Technology Conference,* 2002, pp. 315-320 vol.1.
[23] H. Qiu, J. Lee, J. Lin, and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics," *Journal of Sound and Vibration,* vol. 289, pp. 1066-1090, 2006.
[24] P. Indyk and R. Motwani, "Approximate nearest neighbors: towards removing the curse of dimensionality," in *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, Dallas, Texas, United States, 1998.
[25] V. Sotiris and M. Pecht, "Support Vector Prognostics Analysis of Electronic Products and Systems," in *AAAI Fall Symposium on Artificial Intelligence for Prognostics*, 2007, pp. 120-127.
[26] V. Vapnik, E. Levin, and Y. L. Cun, "Measuring the VC-dimension of a learning machine," *Neural Comput.,* vol. 6, pp. 851-876, 1994.
[27] C. W. Hsu, C. C. Chang, and C. J. Lin, *A practical guide to Support Vector Classification*: Technical report, Department of Computer Science and Information Engineering, National Taiwan University, Taipei., 2003.
[28] H. Ocak and K. A. Loparo, "Estimation of the running speed and bearing defect frequencies of an induction motor from vibration data," *Mechanical Systems and Signal Processing,* vol. 18, pp. 515-533, 2004.