The Development of A Novel Fault Identification Technique by Combining Minimum-Distance Pattern-Recognition and Discrete Wavelet Transform

Pulung Nurprasetio[1], **Komang Bagiasna**[2], **Djoko Suharto**[3], and **Tegoeh Tjahjowidodo**[4]

[1]Associate Professor, [2,3]Professor
Mechanical Engineering Department, Faculty of Mechanical and Aerospace Engineering
Institut Teknologi Bandung, Bandung, Indonesia
e-mail: ipn@labsurya.ms.itb.ac.id (corresponding author)

[4]Assistant Professor, Division of Mechatronics and Design, School of Mechanical and Aerospace Eng.
College of Engineering, Nanyang Technological University, Singapore

Abstract

In this paper, a new idea in machinery fault diagnostic by combining minimum-distance pattern-recognition and discrete wavelet transform (DWT) is presented. The method starts by forming the feature vector, which is basically the position vector of a point in the hyperspace spanned by the set of wavelet coefficients. These coefficients are obtained by expanding the measured signal from faulty machines using DWT. Fault identification is performed based on the distance between the feature vector of the measured signal and the reference feature vector. The basic idea comes from the fact that the wavelet expansion coefficients are unique for a particular signal, considering completeness of the wavelet series. In order to show the effectiveness of the method, two standard machinery faults, namely the unbalance case and the mechanical looseness, were simulated. Experimental results are very promising, since the method successfully identifies the simulated faults.

1 Introduction

Condition Based Maintenance or CBM is a maintenance program in which the decision to repair or recondition is made based on the information obtained from machine condition monitoring [1 – 3]. The standard CBM procedure consists of three steps, the first is monitoring data acquisition, followed by data processing and analysis, before finally one may issue repair decision based on the previous steps. Since the machine is continuously monitored, any fatality may be circumvented. In addition, early diagnostic allows the maintenance department to carefully plan just-in-time actions. As a result, the utilization of personnel, machine, and component will be optimum, and down-time and cost may be retained at its minimum level.

There are two main activities in CBM, i.e., diagnostic and prognostic [1]. Typical activities in diagnostic are fault detection, fault isolation, and fault identification. On the other hand, prognostic deals with prediction before any failure may occur. Despite the difference in activities, both diagnostic and prognostic employ the same CBM techniques.

The focus of this research work is on diagnostic technique development. As for novelty, the paper would like to embark on the use of discrete wavelet transform (DWT) [4]. The idea of employing wavelet transform in diagnostic is relatively new [5]. In 2006, Ren et al. [6] simply take advantage of the denoising power of DWT to purify measured signal before analyzing the data using the standard time domain observation and frequency domain methods. Another interesting effort developed by Goumas et al. [7] utilizes the approximation and detail coefficients of wavelet as interim signal. The augmented vector consisting of the wavelet coefficients are extracted and post processed. The calculated correlation sequence, mean value, and variance are formed into a feature vector. In the end, with the aid of Euclidean, Mahalanobis, Bayes, and Karhunen-Loeve transformation of the covariance matrix, diagnostic of faulty washing machine is performed. Other researchers, Lou and Loparo [8], utilized measured signal decomposition obtained through DWT. The developed diagnostic method is based on the Euclidean distance which is formed from the difference of the variance of the decomposed signals. The last two references belong to the class of coefficient-based method [5]. Recently, Zhu et al. [9] discussed the implementation of continuous wavelet transform (CWT) in fault diagnostic of faulty gear box and bearing. In this case, the method combines CWT and Kolmogorov-Smirnov test.

In this paper, an alternate diagnostic method is developed based on the combination of DWT and minimum-distance pattern-recognition. The feature vector that is used for fault identification will be formed based on the DWT expansion coefficients. This feature vector, which is basically a position vector in the hyperspace spanned by the DWT expansion coefficients, is selected based on the appropriate bandwidth of the measured signal. To identify the fault, the method finds the distance between the feature vectors to the reference position vector in the data base. The method offers simplicity as compared to the available methods, which is the main advantage of the proposed diagnostic procedure. Since it takes advantage of the wavelet decomposition coefficients, the developed method belongs to the class of coefficient-based.

The paper is organized as follows: the problem statement is presented in Section 2, followed by a brief overview of DWT in Section 3. The proposed diagnostic method is outlined in Section 4, and its implementation in a laboratory set up is elaborated in Section 5. In this case, the method will attempt to identify pure unbalance as opposed to simultaneous unbalance and mechanical looseness faults. Finally, the conclusions of this research work are summarized in Section 6.

2 Problem Statement

Fault identification techniques should be able to perform two basic functions. The first is that they must detect and identify similarity pattern of signals that belongs to the same fault. Secondly, they must recognize this pattern in any new signal to classify it as a representation of a particular fault. The fundamental problem is to find a distance measure $d(\bullet)$ that may be used to describe the closeness of the new signal to the known faults. The distance may mean closeness in some arbitrary, abstract property, and does not have to be in the ordinary Euclidean sense [10]. In addition, the distance $d(\bullet)$ must satisfy the *metric* conditions, exemplified for the case of two objects, a and b, as follows:

$$d(a,b) = d(b,a) \quad \text{(symmetric)} \tag{1a}$$
$$d(a,b) \le d(a,c) + d(b,c) \quad \text{(triangle inequality)} \tag{1b}$$
$$d(a,b) \ge 0 \quad \textit{(non-negative)} \tag{1c}$$
$$d(a,b) = 0 \textit{ if, and only if, } a = b \tag{1d}$$

The fault identification problem considered in this work may be then stated as follows: find a distance measure $d(\bullet)$ that will minimize distances between signals measured from faulty machines to show things that are "close". In addition, one needs to develop the procedure to practically implement the method.

3 Discrete Wavelet Transform (DWT)

Every signal $y(t)$ in $L^2(R)$ signal space, i.e., the space of finite energy signals, $\int_0^\infty |y(t)|^2\,dt < \infty$, may be expanded using wavelet series as:

$$y(t) = \sum_j \sum_k d_j(k)\psi_{j,k}(t) \tag{2}$$

in which $\psi_{j,k}(t)$ represents the orthonormal wavelet series [11]. The set of expansion coefficients $d_j(k)$ are called the DWT of $y(t)$. These coefficients may be calculated using the following inner product:

$$d_j(k) = \int_0^\infty y(t)\,\psi_{j,k}(t)\,dt \tag{3}$$

The wavelet expansion (1) is known to converge uniformly [12]. Wavelet series are generated from "mother wavelet' according to the following relation:

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \qquad j,k \in Z \tag{4}$$

with Z: the set of integer values. Mother wavelets are finite duration signals having specific forms and its generation follows specific formulation [12, 13].

The set of wavelet series consists of two main ingredients, namely the scaling functions $\varphi_k(t)$ and the wavelet functions $\psi_{j,k}(t)$. Scaling functions gives DWT the multiresolution aspect. In a simple way, resolution is inversely proportional to frequency. It may be shown that the scaling and wavelet functions spanned $L^2(R)$ in a nested manner. Using the above functions, a more general expression of the series expansion (2) may be given by [11]:

$$y(t) = \sum_k c_j(k)\varphi_{j,k}(t) + \sum_k \sum_j d_j(k)\psi_{j,k}(t) \tag{5}$$

in which $\varphi_{j,k}(t) = 2^{j/2}\varphi(2^j t - k)$, $\psi_{j,k}(t)$ as in (4), and $j,k \in Z$, the set of all integers. The scaling and wavelet functions are orthogonal, and both are derived from the same family of mother wavelet. Furthermore, the scaling or approximation coefficients, $c_j(k)$, and the wavelet or detail coefficients, $d_j(k)$ may be obtained, respectively, as:

$$\begin{aligned} c_j(k) &= \sum_m h(m-2k)c_{j+1}(m) \\ d_j(k) &= \sum_m h_1(m-2k)c_{j+1}(m) \end{aligned} \tag{6}$$

The above relations are known to be the filtering and down-sampling or decimating equations. The filter $h(k)$ is a low-pass filter while $h_1(k)$ is a high-pass and both are of finite impulse response (FIR) type. The coefficients of the filter $h(k)$ and $h_1(k)$ basically are the weighting factors of scaling and wavelet function generators, respectively [11]. For a finite number of expansion terms, equation (5) may be rewritten as:

$$y(t) = \sum_k c_{j0}(k)\varphi_{j0,k}(t) + \sum_k \sum_{j=j0}^{J+j0-1} d_j(k)\psi_{j,k}(t) \tag{7}$$

in which J is the total number of decomposition levels, j_0 denotes the lowest frequency band or level, and $J+j_0-1$ is the highest frequency level. By equation (6), the expansion or decomposition coefficients are obtained starting from the highest level. Due to this feature, in numerical soft wares, the decomposed signals are normally designated as $D_j(t)$ and $A_j(t)$, in which the D stands for 'detail', A stands for 'approximation', and j is the decomposition level [4, 14]. Using the above signals, equation (7) may be rewritten as:

$$y(t) = A_J(t) + \sum_{j=J}^{1} D_j(t) \tag{8}$$

where j runs from 1 to J, $A_J(t) = \sum_k c_{j0}(k)\phi_{j0,k}(t)$,

and $D_{J-j+1}(t) = \sum_k d_{j0+j-1}(k)\psi_{j0+j-1,k}(t)$.

4 Combination of Minimum-Distance Pattern-Recognition and DWT

As has been explained in Section 2, in order to implement the minimum-distance pattern-recognition technique, one needs to define a distance measure. Since the DWT produces expansion coefficients in the form of the set $\{c_j(k), d_j(k)\}$, and the set is unique, a natural choice for the feature vector will be the position vectors defined in the hyperspace spanned by the approximation and detail coefficients. These position vectors may thus be written as:

$$\theta_o = [c_j(k), d_j(k)]^T \tag{9}$$

For a particular machinery fault, a reference feature vector may be set up based on measurements performed on machines with known faults. Fault identification will be determined based on the following distance measure:

$$d = \left\| \theta - \theta_o \right\| \tag{10}$$

In the above equation, $\|\bullet\|$ denotes l_2-norm, θ is the feature vector of an unidentified fault, and θ_o is the reference feature vector. The set of approximation and detail coefficients are selected based on the theoretical bandwidth of the known or simulated faults. This will allow the user to take advantage of the decomposition capability of DWT.

The standard procedure of performing fault identification using the proposed method may thus be written as follows [15]:

1. Perform measurement on running faulty machines with known faults and obtain the proper signals
2. Obtain the set of approximation ($c_j(k)$'s) and detail coefficients ($d_j(k)$'s) for the appropriate bandwidth using DWT and form the reference feature vector, θ_o's. Repeat this step for each individual faults.
3. Get measured signal(s) from similar machines with unknown or unidentified fault, and obtain the set of approximation ($c_j(k)$'s) and detail coefficients ($d_j(k)$'s) for the appropriate bandwidth using DWT and form its feature vector, θ.
4. Calculate the distances between the feature vector as a result of Step 3 and the reference feature vectors from Step 2 using equation (10).
5. Identify the fault based on the smallest distance obtained from Step 4.

5 Laboratory Experiment

In order to show the effectiveness of the method, a laboratory experiment was performed using the available vibration test rig [16]. The rig consisted of a continuous beam with pin-pin supports and an unbalanced-mass exciter as depicted in Figure 1.

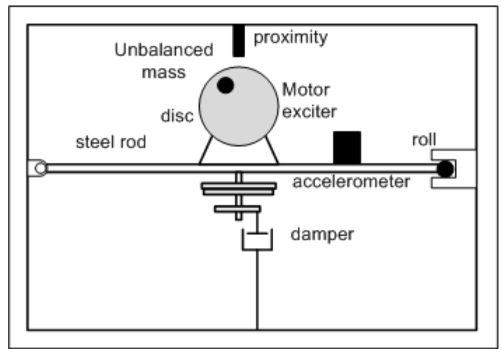

Figure 1: Experimental test rig

The pin support at the left end is fixed, while the one at the right end may be replaced using a tight roller support or a slack support to simulate mechanical-looseness. The measurement was conducted using Kyowa AS-2C CD-4659 piezoresistive accelerometer. The accelerometer was cross-calibrated using Lutron VB-8202 Vibration Meter. As a key phasor, Autonics PR12-4DN is employed. Data acquisition was carried out by NI 800820A-01 external USB card, and the data acquisition was governed by LabVIEW SignalExpress which was set at 1 kHz sampling frequency.

As for the experiment, two cases were simulated, namely the pure unbalance case, and the combination of unbalance and mechanical looseness. For the first case, the beam was supported tightly using pin-type supports at both ends, while for the second case, the housing of the right roller was replaced, resulting in a loose support.

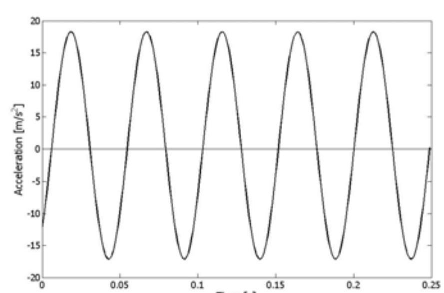

Figure 2: Acceleration signal of the unbalance case

Typical measurement signals in the time domain are depicted in Figure 2 and 3, for the unbalance and the combination of unbalance and mechanical looseness, respectively [17].

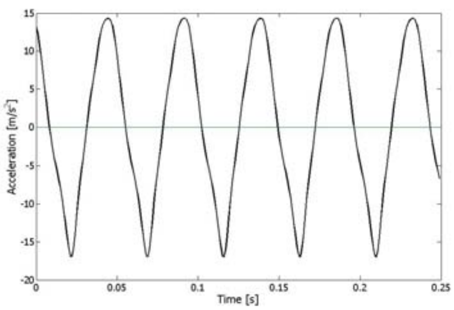

Figure 3: Acceleration signal of the unbalance plus mechanical looseness case

Typical frequency profiles of the simulated cases are shown in Figures 4 and 5, expressed in terms of the periodogram of the measured signals. Figure 4 is for the unbalanced case, and Figure 5 is for the combination of unbalanced and mechanical looseness case.

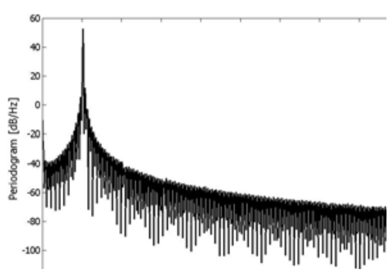

Figure 4: Periodogram of the acceleration signal for the unbalance case

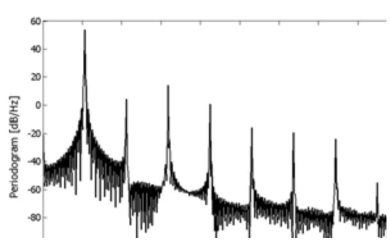

Figure 5: Periodogram of the acceleration signal for the unbalance plus mechanical looseness case

From the periodogram, it is apparent that the frequency profiles of the simulated cases are quite distinct. Since the objective of the simulation is to show that the method is indeed effective in detecting faults, the simulation is an ideal situation, despite the fact that it is a bit trivial. Using the procedure outlined in Section 4, the distance measurements were performed, and the complete results are presented in Table 1. In the DWT calculation, J = 4and $j_0 = 0$. Due to the distinct features of the theoretical fault frequency profiles, only $\{c_{j0}(k)\}$ and $\{d_{j0}(k)\}$ are used to form the feature vectors.

Table 1. Distance Measurement Results

	Data Signal	
Data Signal	Unbalance	Unbalance + looseness
Unbalance	0	212.8
Unbalance + looseness	212.8	0
Test #1	7.0	213.8
Test #2	187.7	22.8

The explanation of the figures in Table 1 is as follows:

- The first and second rows denote the distances between the reference feature vectors for the unbalance and combination of unbalance and looseness. Obviously, the auto-distance, i.e., numbers in the diagonal positions, will be zero, and the cross-distance, i.e., the off-diagonal positions, will be large and symmetric. This fact shows that the distance formulation satisfies the metric requirements as depicted in equation (1).
- The third row gives the distance between the test signal, taken from the unbalance case but different than the set used in forming the reference feature vector, and the reference feature vectors for both simulated cases. In this case, the distances between signals taken from the same fault are small, and the distance between signals taken from different faults are large. This shows that the method is effective in detecting the simulated faults which are assumed to be unknown in the distance calculations.
- The fourth row basically is similar to the third row, except that the test signal is taken from the combination of the unbalance and mechanical looseness but different than the set used in forming the reference feature vectors. Again, in this case the results confirm the effectiveness of the proposed method in detecting unidentified faults.

6 Conclusions

This paper presents a novel method in machinery fault identification based on the combination of minimum-distance pattern-recognition technique and DWT. The feature vectors are formed based on the DWT decomposition, i.e., approximation and detail, coefficients.

Experimental validation showed promising results, which confirms the effectiveness of the method.

Despite the above encouraging results, the robustness issue still needs to be addressed. In addition, the sizes of the feature vectors are relatively large. This inspired the idea of incorporating signal compression methods in the future. A viable path will be to implement time-series modeling to the expansion coefficients and compose the smaller feature vectors from the time-series model coefficients.

Acknowledgements

The authors gratefully acknowledged the assist of Mr. Hilarius Tutut Sandewan, who verified the method through numerical simulations, and Mr. Mirza Bamanto and Mr. Aris Ilmi who helped to set up the test rig and performed the laboratory experiments.

References

[1] Jardine, A.K.S., Lin, D., and Banjevic, D., "A review on machinery diagnostics and prognostics condition-based maintenance", *Mechanical Systems and Signal Processing* 20 (2006) 1483-1510.

[2] Scheffer, C., and Girdhar, P., *Practical Machinery Vibration Analysis and Predictive Maintenance*, Newnes, Oxford, 2004

[3] Braun, S., *Mechanical Signature Analysis – Theory and Aplication*, Academic Press, London, 1986

[4] Misiti, M., Misiti, Y., Oppenheim, G., and Poggi, J-M., *Wavelet Toolbox For Use with MATLAB®*, The MathWorks, 1997

[5] Peng, Z.K., and Chu, F.L., "Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography", *Mechanical Systems and Signal Processing*, 18 (2004) 199-221

[6] Ren, X.P., Shao, W., Yang, W.Z., and Su, F.Q., "Extraction of Gearbox Fault Features from Vibration Signal using Wavelet Transform", *Journal of Physics*: Conference Series, Vol. 48, pp. 490 – 494, 2006

[7] Goumas, S.K., Zervakis, M.E., and Stavrakakis, G.S., "Classification of Washing Machines Vibration Signals using Discrete Wavelet Analysis for Feature Extraction", *IEEE Transactions on Instrumentation and Measurement*, Vol. 51, No. 3, June 2002, pp. 497 – 508

[8] Lou, X., and Loparo, K.L., "Bearing Fault Diagnosis Based on Wavelet Transform and Fuzzy Inference", *Mechanical Systems and Signal Processing*, Vol. 18, 2004, pp. 1077 – 1095

[9] Zhu, Z.K., Yan, R., Luo, L., Feng, Z.H, and Kong, F.R., "Detection of signal transients based on wavelet and statistics for machine fault diagnosis", *Mechanical Systems and Signal Processing*, Vol. 23, 2009, pp. 1076 – 1097

[10] Sebestyen, G.S., *Decision-Making Processes in Pattern Recognition*, The Maxmillan Company, New York, 1962

[11] Burrus, C.S., Gopinath, R.A., and Guo, H., *Introduction to Wavelets and Wavelet Transforms – A Primer*, Prentice-Hall, 1998

[12] Hess-Nielsen, N., and Wickerhauser, M.V., "Wavelets and Time-Frequency Analysis", *Proceedings of the IEEE*, Vol. 84, No. 4, April 1996, pp. 523 – 540

[13] Daubechies, I., *Ten Lectures on Wavelets*, CMBS-NSF Series in Applied Mathematics, SIAM, 1992

[14] Fugal, D.L., *Conceptual Wavelets in Digital Signal Processing*, Space & Signals Technologies, 2009

[15] Sandewan, H.T., *The Development of a Diagnostic Technique by Combining Minimum-Distance Pattern-Recognition and Discrete Wavelet Transform*, Master's Thesis, Mech. Eng. Dept., Faculty of Mechanical and Aerospace Engineering, ITB, Bandung, 2010 (in Indonesian).

[16] Bamanto, M., *Vibration Test Rig Modification to Simulate Unbalance and Mechanical Looseness Simultaneously*, Undergraduate Final Project Report, Mech. Eng. Dept., Faculty of Mechanical and Aerospace Engineering, ITB, Bandung, 2010 (in Indonesian).

[17] Irmi, A., *The Implementation of a Novel Diagnostic Technique by Combining Minimum-Distance Pattern Recognition and Discrete Wavelet Transform in Vibration Testing*, Undergraduate Final Project Report, Mech. Eng. Dept., Faculty of Mechanical and Aerospace Engineering, ITB, Bandung, 2010 (in Indonesian).