A Concept for Intelligent Fault Detection Using a Multi-resolution Analysis

Ousmane Diallo, Dimitri Mavris, Ph.D.
Aerospace Systems Design Laboratory
Georgia Institute of Technology
Atlanta, GA, USA
ousmane.diallo@gatech.edu

Abstract - *As a consequence of recent deregulation in the electrical power production industry, many new private entrepreneurs with no prior experience in power plant operation have entered into the power plant business. To hedge their business risks, they outsource operation and maintenance activities to third party service providers with whom they share risks/rewards of plant performance. With the cost benefit of preventing compressor failure estimated at ten million dollars or more, techniques for detecting faults are important. In this paper, a systematic process is proposed for intelligent failure precursor detection. The wavelet packet transform is used to perform a multi-resolution analysis of health monitoring sensor data to extract their signal features. Then the probabilistic principal component analysis is utilized to fuse them into a few uncorrelated variables. Finally, a Bayesian hypothesis testing method is applied for anomaly detection. The proposed process is successfully applied to a gas turbine compressor failure as an example.*

Keywords: diagnostics, feature extraction, intelligent fault detection, wavelet packet, data fusion, gas turbine.

1 Introduction

As a consequence of the recent deregulation in the electrical power production industry, there has been a shift in the traditional ownership of power plants and the way they are operated. Many new private entrepreneurs with no prior experience in power plant operation have invested into the power plant business. Thus, to hedge their business risks, private entrepreneurs enter into long-term service agreement (LTSA) with third parties for their operation and maintenance (O&M) activities. Thereby, the original equipment manufacturers (OEMs) become the natural choices as third party O&M providers because they know and understand their designed products best and will be willing to guarantee their operations. Together, the major OEMs (e.g., General Electric (GE) Energy, Siemens Power Generation, Alstom Power, Mitsubishi Heavy Industries, and Ansaldo Energia) represented about 94% of the global market for the period of 2000 through 2004 [1]. Each of these main gas turbine OEMs has its own set of definitions and foreseeable benefits to the plant owners of their LTSA offerings. Thus, the major OEMs have invested huge amounts of money to develop preventive maintenance strategies to minimize the occurrence of the normally costly unplanned outages resulting from failures of equipment covered under LTSA contracts.

The high potential for cost benefits to gas turbine OEMs when failures can be prevented raises the importance of techniques for detecting faults in gas turbines. In this paper, a systematic process is proposed for intelligent detection of failure precursory events. Section 2 sets the context and background regarding power plant O&M and the background for the problem addressed. Section 3 presents the steps of the proposed approach to detect catastrophic failure precursors. Then an illustrative example of application to a gas turbine compressor failure problem is presented in Section 4 followed by a brief conclusion in Section 5.

2 Power plant O&M background

Typically, the LTSA contracts work like insurance policies where the manufacturer guarantees a given level of power output and/or efficiency over several years. They also provide repair, replacement, and upgrade parts to the degrading power plant. Overall, it is supposed to be a win-win partnership for both parties, wherein they share the operational risks as well as the rewards of extra power plant performance. Besides the apparent benefits for sharing the plant operational risks with the third party O&M provider, there are other benefits both for the plant owners and the OEM as LTSA provider. The LTSA contract raises the plant re-sale value while, for the OEM, the equipment under contract provides unprecedented access to "a live laboratory" that should allow the OEM to learn from eventual design shortcomings of previous gas turbine designs in order to improve upon future designs, ultimately giving them a competitive advantage.

2.1 Power plant operation and maintenance

Currently, the O&M expenditures of a typical power plant are an important part of the total life cycle cost consisting of 15% to 20%, while equipment maintenance costs account for approximately 10% to 15% [2]. There is always a cost associated with an outage whether it is planned or unplanned. Thus, to make the LTSA contracts profitable, the providers need to reduce the number of unplanned outages because the consequence of such unplanned outages can be expensive. Typically under a LTSA contract, the provider has to pay the plant owners a liquidated damage for each forced outage. In general, the liquidated damage cost for a forced outage includes: the loss of production cost, the repair cost, the cost to buy power to meet the quantity that the forced outage plant was dispatched for at usually higher prices, and eventual penalties.

2.2 Problem Background

With the steep cost of potential liquidated damage amounting to multi-millions of dollars associated with not meeting the reliability and durability requirement, LTSA providers need to develop strategies so that the revenues from the contracts exceed the cost of the involved risks. In fact, according to a report of the Electric Power Research Institute (EPRI), the cost benefit from preventing a General Electric gas turbine 7FA and 9FA technology compressor failure is estimated to be ten to twenty million dollars [3].

It is well known in the heavy-gas turbine industry that over the years the profit margin on sale of new gas turbines has been shrinking for OEMs while the LTSA, for the most part, assured OEMs of the sales from upgrading new components to sustain degrading gas turbine performance throughout the duration of the LTSA. If managed well, the LTSA market can be very lucrative. Thus, OEMs have been investing huge amounts of money to develop strategies to avoid unplanned plant outages in the first place. For example, OEMs like GE Energy created a Power Answer Center in Atlanta, GA, where all power plants under its LTSA contract are continuously monitored using installed sensors; the data are recorded and stored for post processing to detect any abnormal trends. The illustrative Figure 1 shows the GE Power Answer architecture wherein the on-site monitor compares the actual unit performance with baseline predictions and provides the first level of anomaly detection and notification.

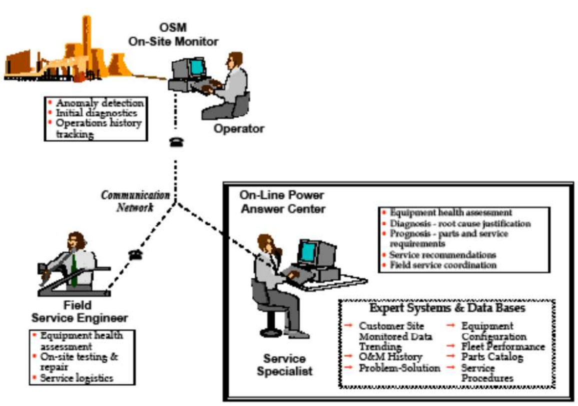

Figure 1. GE Monitoring & Diagnostics concept [1].

Major OEMs like GE have the ability to monitor hundreds of units throughout the world in real time in order to establish knowledge to detect faults before they can develop into failure. This is both challenging and can yield some advantages toward sustaining the technological competitive advantage of an OEM in the long run. However, some of the challenges faced by the most important OEMs are due to the fact that they may have hundreds of units throughout the world that differ with respect to machine-to-machine variation, uncertainties of machine degradation, and power plant operating conditions. All of these distinctions make it difficult to recognize some of the rare or short lasting fault events that might lead to engine trips. Thus, preventive maintenance planners have to create rules representing all possible detections for a given set of gas turbines. Despite all of the effort to avoid forced outages, there are still undetected failure precursors that led to catastrophic failure as reported by EPRI in its 2007 updated report [4].

3 Detection of failure precursors

Though main OEMs have improved their ability to monitor and store gas turbine condition data in real time, the data analysis capability has been lagging for the most part. In recent years, there have been new and improved techniques such as condition-based monitoring (CBM) to help detect anomalies in their early stages of development. Currently, the new techniques have not allowed to totally resolve the issue of missed detections of all the anomalies. Although their merit is well accepted, their practical implementation is still inefficient because these techniques tend to be theoretical, difficult, and/or expensive to apply to real world problems. Therefore, the method proposed herein intends to take advantage of the monitoring sensors to capture catastrophic failure precursors.

In general, the health and condition of power plants are monitored using two types of sensors: the static or process-related sensors used for temperature, pressure, and flow rate measurements, and the sensors characterized by their high-bandwidth used for high-frequency signals like the vibration measurements. There are many time-frequency techniques reported in the research literature such as the Wigner-Ville distribution, the Choi-Williams distribution, the Fourier transform and its variants (e.g., short time Fourier transforms), etc; but the wavelet transform is the best one to deal with short lasting anomalies and sharp discontinuities [5]. The following subsections provide a brief overview of the wavelet transform followed by presentation of by a step-by-step explanation of the proposed approach.

3.1 Wavelet transforms overview

The time-frequency analysis techniques are appropriate when dealing with time series because more information can be extracted about small variations of a signal in the combination of time and frequency domains than can be extracted in the time domain alone. The Fourier transform is the most popular time-frequency domain analysis technique because of its ability to decompose an energy limited signal *f(t)* so as to analyze the signal in the time domain for its frequency contents *F(ω)* as defined by Equations 1 and 2:

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\varpi)e^{i\omega t}d\omega \tag{1}$$

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t}dt \tag{2}$$

However, the Fourier transform provides only the global information on the frequencies of a signal, it cannot provide local information if the spectral composition of a signal changes rapidly with time [6]. In other words, once a signal is Fourier transformed, all the time domain information is lost, while the wavelet transform conserves both the time and the frequency information. Thus, the wavelet transform is an improvement over Fourier transforms for time-frequency analysis. Wavelet transforms decompose a given signal through two filters: a low-pass filter that provides a low frequency part which trends and smoothes the original signal (i.e., approximation), and a high-pass filter that provides the high frequency part (i.e., details) which reveals local properties such as anomalies.

3.1.1 Mathematical overview of Wavelet Transforms

There is a panoply of literature on the theory of wavelet transforms and its applications [7, 8]. Just like the Fourier transforms, the wavelet transform can be defined for any square-integrable function $L^2(\Re)$ [9]. But instead of using the harmonics, $e^{i\omega t}$, the wavelet basis, ψ, called a mother wavelet function, is used and defined as:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \tag{3}$$

where a is the dilation or scaling parameter and b is the time location or translation parameter. Thus the wavelet transform of a signal *f(t)* is computed as follows [8]:

$$W_f(a,b) = \int_{-\infty}^{\infty} f(t)\psi_{a,b}(t)dt \tag{4}$$

3.1.2 Wavelet Packets

The standard wavelet transform has a limitation because it can only decompose the low-frequency part of a signal. The wavelet packet transform was introduced to resolve that limitation; it has the ability to decompose both the approximation part as well as the detail part. Thus, the wavelet packet transform decomposes a signal into more detailed components than the standard wavelet transform could, thereby yielding more information about the signal. For that reason, it is more advantageous to use the wavelet packet transform to realize the multi-resolution analysis (MRA) by decomposing both the low frequency and high frequency components of a signal into subspaces so as to obtain finer and adjustable resolution [10]. Figure 2 illustrates a wavelet packet decomposition of a signal, s.

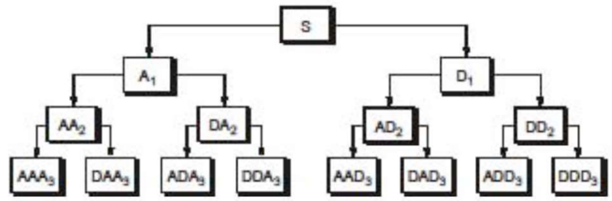

Figure 2. Wavelet packet decomposition [11]; signal s is decomposed as s = A1+ +AAD3+DAD3+ADD3+DDD3 at level 3.

3.2 Steps for failure precursor detection

As mentioned above, the proposed approach intends to take advantage of monitoring sensors to capture catastrophic failure precursors. Figure 3 shows a flowchart of the proposed methodology for intelligent failure precursor detection using multi-resolution analysis. A step-by-step explanation of each block in the flowchart is presented in the subsections below.

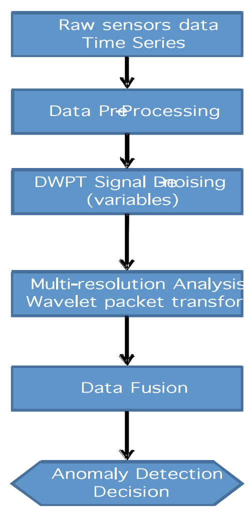

Figure 3. Intelligent failure precursor detection.

3.2.1 Raw time series data collection

The systems health and operating condition parameters are continuously monitored and collected using installed sensors and stored for potential post-processing. The installed sensors for heavy-duty gas turbines typically include the two types mentioned earlier, static or process-related sensors (used for pressure, temperature, and flow rate measurements) and high-bandwidth sensors used to measure high-frequency measurements (e.g., vibration measurement).

3.2.2 Data Pre-Processing

The pre-processing of the raw data is a necessary step for a couple of important reasons. First, the OEMs will not want to share their proprietary data on equipment malfunctioning because that may affect their competitive advantages. Secondly, the sensors monitor different health parameters (e.g., temperature, pressure, vibration, etc) that are recorded in different units and orders of magnitude. For instance, a typical normal base load operation of GE's 7FA+e gas turbine technology can have a compressor discharge temperature measurement in the range of 600 to 800 degrees Fahrenheit, while the vibration sensor measurements could be on the order of 1/10 of an inch per second. Therefore, an analysis with the raw measurement could be artificially skewed towards the variables with higher absolute values. Thus, the pre-processing step consists of normalizing each measured parameter value by the mean value of that variable measurement, and eliminating the visual outliers that would misrepresent the finding and affect the accuracy of the conclusion.

3.2.3 DWPT signal de-noising

The de-noising step is essential because a sensor measurement signal is always tainted by noise. In [12] the authors presented a de-noising technique that adequately removes the noise by combining the discrete wavelet packet transform (DWPT) and Bayesian thresholding. The result is the removal of just the noise without the drawbacks of many other de-noising algorithms that either remove useful information along with the noise or remove too little noise thus leaving some noise in the signal.

3.2.4 Multi-resolution analysis using discrete wavelet packet decomposition

In this step, the de-noised signal is decomposed at an appropriate level of resolution (as in Figure 2) to get the approximation and the detail components. The content of each component resulting from the decomposition can be analyzed. Once the decomposed tree is obtained, the energy content of the scaling function (approximation) and the wavelet functions (details) representing the nodes of the tree is calculated as:

$$E_{j,n} = \int_{-\infty}^{\infty} \left(W_{j,n,k}(dt)\right)^2 dt = \sum_k W_{j,n,k}^2 \tag{5}$$

where $W_{j,n,k}$ is the wavelet packet transform coefficient, j is the level, k is the translation, and n is the modulation parameter (approximation or detail). The energy content of each node will then be used as the signal features.

3.2.5 Data Fusion using PPCA

The goal of the data fusion step is to combine pieces of information from a potentially correlated multi-sensory data set system into fewer uncorrelated variables that allow for drawing a more adequate conclusion than one could get from each individual sensor. Thus, the probabilistic principal component analysis (PPCA) is used to merge the information from the sensors of interest. To perform the PPCA, the steps of the principal components analysis are executed, then the notion of maximum likelihood and the variance of the reduced data is calculated using matrices, where only the most significant weights obtained from the standard PCA are used as entries in a maximum likelihood matrix. The PPCA is an improved version of the standard PCA as it has the advantage of taking into account data uncertainty [13].

3.2.6 Anomaly Detection decision

The Bayesian evaluation method is applied to the following error-related metrics for hypothesis testing:

$$\Delta(i) = r_*(i) - r(i) \tag{6}$$

where $r_*(i)$ is the true measurement and $r(i)$ is the estimation value. Specifically, the hypothesis is applied to the weighted error ε_w, which is the sum of each of the weighted errors, ε_{wi}, of the i^{th} principal component as defined by the following equations:

$$\varepsilon_w = \sum_{i-1}^{M} \varepsilon_{wi} \tag{7}$$

$$\varepsilon_{wi} = C_i * \left(Y_{i\,original} - Y_{i\,model}\right) \tag{8}$$

where M is the number of principal components, C_i is the contribution of information contained in the i^{th} principal component (PC_i), $Y_{ioriginal}$ is the i^{th} probabilistic principal component of the original signal, and Y_{imodel} is the i^{th} probabilistic principal component of the model signal. Thus, the Bayesian evaluation method for hypothesis testing is conducted with a binary outcome. The anomaly function is defined as $H(t)$, which is the vector of the Bayesian hypothesis testing result with null and alternative hypotheses defined as follows:

null hypothesis H0: $|\varepsilon_{wi}| \le \varepsilon$, $H(t) = 1$

alternative hypothesis Ha: $|\varepsilon_{wi}| > \varepsilon$, $H(t) = 0$

where ε is the pre-determined threshold value. Then the function $H(t)$, which has values of 1 or 0, can be plotted over time. Thus, a $H(t)$ value of 1 is a healthy state and a $H(t)$ value of 0 is an anomaly. Therefore, the appearance of the first value of $H(t) = 0$ can be considered as the first failure precursor.

4 Illustration Example

As an illustrative example, the proposed methodology is applied to a gas turbine compressor failure problem. The test unit is a gas turbine compressor with 8 health monitoring sensors as summarized in Table 1. In this example, the gas turbine compressor failed on June 24, 2006 at 18:18. The gas turbine manufacturer found through a post compressor failure analysis that there was a failure precursor event on June 20th at 23:30 that was previously missed.

Table 1. Gas turbine health monitoring sensors.

Sensors	Description
X1	Compressor health parameter 1
X2	Compressor health parameter 2
X3	Inlet guide vane (operating condition)
X4	Gas turbine output (system condition)
X5	Compressor seismic vibration 1
X6	Compressor seismic vibration 2
X7	Turbine seismic vibration 3
X8	Comp. inlet temp.(operation condition)

4.1 Process steps

Step 1:
A query is made to the stored database to retrieve measurements for the 8 sensors of interest at 5-second intervals from June 19, 2006 at 00:00 to June 25, 2006 at 00:00.

Step 2:
The raw data is normalized using the mean value of each variable (sensor). The normalized sensor readings are within the same order of magnitude with a mean value of 1 for each variable.

Step 3:
All normalized raw sensor data is de-noised using DWPTs. For example, Figure 4 shows the removed noise from sensor X1.

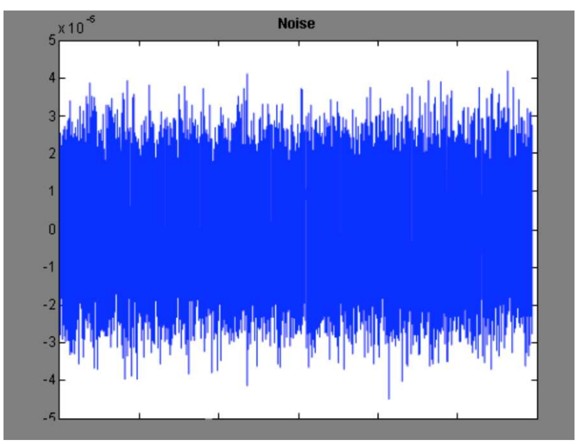

Figure 4. Noise removed from signal X1 using DWPT.

Step 4:
Each variable signal is decomposed into a 3-level tree (Figure 5) using the DWPT and the "Daubechies 4" wavelet mother function [7]. The energy content of each of the 8 nodes representing each wavelet component of the level 3 is calculated and will serve as the signal feature characteristic. It turns out that each of the 8 sensors has over 99.9% of its energy content at the approximation node, which is node 7 or (3, 0) in Figure 5. Therefore, the approximation will be used as a representative of the actual signal in the subsequent steps.

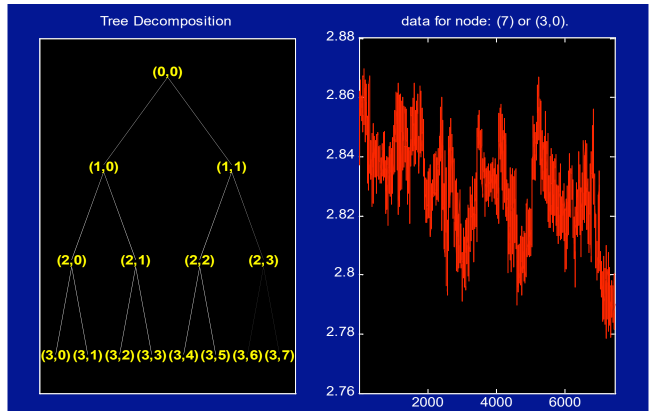

Figure 5. Tree decomposition of signal of variable X1.

Step 5:
First, the standard PCA steps are executed to determine the principal components (PC). As shown on Figure 6, to maintain at least 95% of the original information in the model (i.e., 95% confidence level), the first 3 PCs representing 99.326% of the original information should be retained. Therefore, the eigenvectors corresponding to the retained eigenvalues represent the principal components and are shown in Table 2.

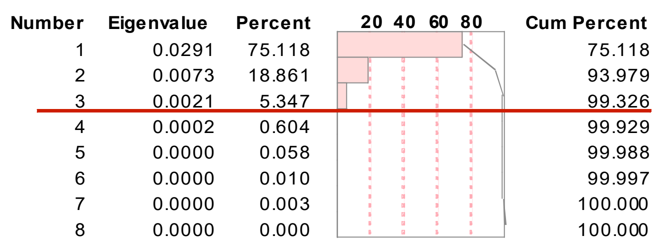

Figure 6. Pareto chart of eigenvalues contribution.

Table 2. Principal components.

	3 Principal Components for 99.3%		
Variables	PC1 (75.1%)	PC2 (18.9%)	PC3 (5.3%)
X1	-0.02085	0.01348	-0.00162
X2	0.0055	-0.01277	0.00362
X3	0.00016	-0.00101	-0.00009
X4	-0.03429	0.03691	-0.00244
X4	0.59281	0.7203	-0.35998
X6	0.46975	0.05324	0.88042
X7	0.65235	-0.69038	-0.30792
X8	0.02671	-0.00045	0.02109

Then the PPCA parameters are calculated with the maximum likelihood weight matrix first by setting to 0 any PC weight that is less than 0.1, as well as other parameters (e.g., the isentropic noise covariance, the prediction error unique to response, the data matrix and the variance of reduced dimension).

Step 6:
Finally, by setting the threshold at $\varepsilon = 0.095$, the result of the Bayesian hypothesis testing is the binary function $H(t)$ with entry values of "1" and "0" obtained and shown in Figure 7. There are four defects, with two of them minutes from each other on June 20, 2006 at 23:30.

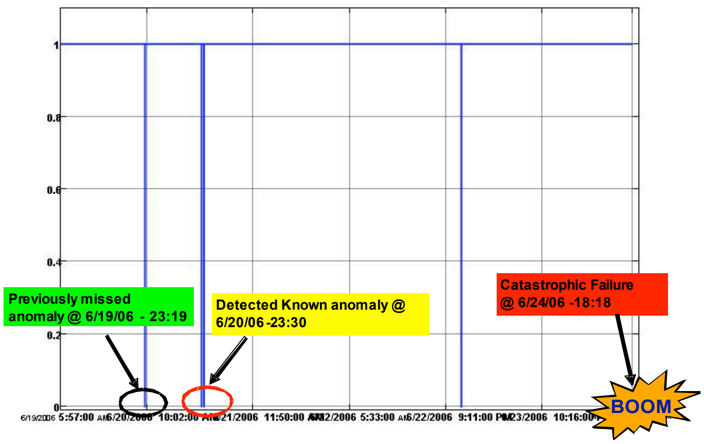

Figure 7. Anomaly detection.

5 Conclusions

The LTSA market can be a very lucrative if the level of liquidated damage is managed to a minimum cost. In this paper, a systematic multi-step approach is presented to detect precursory anomalies that might have led to catastrophic gas turbine compressor failure so as to reduce or even eliminate unplanned power plant outages. The proposed approach is promising as it allows the detection of failure precursory anomalies related to potential catastrophic failures that may be anticipated or suspected by the manufacturer. Moreover, the proposed method has been used successfully to detect a failure precursor event that happened on June 19th, 2009 at 23:19, about 24 hours before the manufacturer's first precursor identification. Therefore, the proposed approach is able to detect failure precursors that are missed by the gas turbine manufacturer.

References

[1] H. Thaler, "Global gas turbine markets and trends - What does the future hold?," The Future of Gas Turbine Technology – 3rd International Conference, Brussels, Oct. 2006.

[2] H.G. Stoll, "Creating owner's competitive advantage through contractual services," GE Power Systems, Technical Report, GER-4208, 2001.

[3] Electric Power Research Institute (EPRI), "GE FA compressor dependability (phase 2)," EPRI Report, Palo Alto, CA, Jan. 2008.

[4] Electric Power Research Institute (EPRI), "Gas turbine compressor dependability: Investigation of the MS7001F/9001F R0 compressor stage blade designs and failures–2007 update," EPRI Report, Palo Alto, CA, Mar. 2008.

[5] A. Graps, "An introduction to wavelets," *IEEE computational science & engineering*, Vol. 2, No. 2, pp. 50-61, 1995.

[6] A. Misrikhanov, "Wavelet transform methods: Application in electroenergetics," *Automation and remote control*, Vol. 67, No. 5, pp. 682-697, 2006.

[7] I. Daubechies, *Ten lectures on wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.

[8] C.K. Chui, *An introduction to wavelets*, Academic Press Professional, Inc., San Diego, CA, 1992.

[9] Y. Wu and R. Du, "Feature extraction and assessment using wavelet packets for monitoring of machining processes," *Mechanical systems and signal processing*, Vol. 10, No. 1, pp. 29-53, 1996.

[10] X. Jiang and H. Adeli, "Wavelet packet-autocorrelation function method for traffic flow pattern analysis," *Computer-aided civil and infrastructure engineering*, Vol. 19, No. 5, pp. 324-337, 2004.

[11] The Mathworks, "Wavelet toolbox 4 – User's guide," Matlab 7.8.0 (R2009a). 1999-2009.

[12] X. Jiang, S. Mahadevan and H. Adeli, "Bayesian wavelet packet denoising for structural system identification," *Progress in structural engineering and materials*, Vol. 14, No. 2, pp. 333-356, 2006.

[13] M. Tipping and C. Bishop, "Probabilistic principal component analysis," *Journal of the Royal Statistical Society*. Series B, Statistical Methodology, pp. 611-622, 1999.