AIAA JOURNAL
Vol. 38, No. 6, June 2000

Fusion of Wavelet Packets and Neural Network in Detection of Composites

Yao-Jun Wu,* Xi-Zhi Shi,† and Tian Ge Zhuang‡
Shanghai Jiao Tong University, 200030 Shanghai, People's Republic of China

A new data-fusion method is proposed for the damage detection of anisotropic composite materials. Based on signal processing theory, we combine wavelet packets, which can decompose signals into a tiling of plane of the time frequency and the feature information in different frequency bands, with autoregression spectrum analysis to extract features and recognize the characteristic signals sampled at experiments of damage detection of composites by vibration. These features are fed into the wavelet neural network as the input patterns for training and classifying. Analysis on the signals obtained in the damage detection experiment of composites demonstrates the effectiveness of the proposed method.

Introduction

RECENTLY on-line damage detection of flexible structures has been a subject of intensive investigation. Dynamic response, modal parameters, and spectrum analysis are applied in the field of material damage detection. The candidate models for structures with various types of damage are designated as the patterns. These different patterns are organized into pattern classes according to the location and the severity of the damage. System identification is performed to extract the transfer functions as the features of the structural system. These features are then trained in backpropagation (BP) neural network.[1–3] However, compared with the authors' experiments to extract the feature coefficients of modal and frequency spectrum and to train the patterns for input to the BP neural network, it is difficult to effectively detect the different damage of anisotropic composites with large divergence. Method of extraction of features is therefore an important element. Neural networks are also applied to detect the damage of composites.[4] The magnitude spectrum of signals trains and tests the patterns for input to the neural network. However, it is difficult to draw features with a nonstationary signal from damage detection composites. Moreover, if the neural network is to be kept to a manageable size, compression of true features is an important step. The successful method must be able to analyze nonstationary signals, ignore the noise, and detect relevant features.

Wavelet transform with good time-frequency localization is applied in signal processing. An intuitive concept is the multiresolution signal representation,[5,6] which considers a signal at different levels of resolution. Wavelet decomposition is a successive approximation method that adds more and more projections onto detail spaces spanned by wavelets and their shifts at different scales. Wavelet transform characterizes the fine component of nonstationary signals. The fine component implies high frequency or small scale, and how this changes over a brief duration of time. The wavelet transform divides up the timescale plane in such a way that high-frequency activity is described with very sharp time resolution. This method has been applied to mechanical systems[7–9] but less in composite systems of damage identification.[10–13]

Wavelet packets were first introduced by Coifman and Meyer[14] as a library of orthogonal bases for square-integrable real functions $[L^2(R)]$. Implementation of a best-basis selection procedure for a prescribed signal (or a family of signals) requires the introduction of an appropriate cost function, which translates the selection procedure into a minimization process.[15] Choosing a cost function is intimately related to the specific nature of the application. Entropy, for example, may constitute a reasonable choice if signal compression, identification, or classification are the applications of interest. Statistical analysis of the best-basis coefficients may be used as a signature, representing the original signal. Both the wavelet packet decomposition of Coifman and Wickerhauser[15] and the extended algorithm proposed by Herley et al.,[16] are sensitive to the signal location with respect to the chosen time origin. Signals of specimens in damaged composites may be nonstationary, and their features often occur over several scales. This feature can be used to great advantage for analysis of nonstationary signals. Multiple-scale activity is commonplace in detection and diagnostics of composite materials. The reliability of vibration signals based on wavelet packets analysis demonstrates that the multiresolution property of such transforms is beneficial to the damage detection of composites.

Wavelet neural network has been extensively researched recently,[17–19] and it is a potentially powerful detection system. Because of the abilities to decompose signals into a tiling of the time-frequency plane and to learn and memorize the patterns as an associative memory, special attention has been drawn to the use of wavelet neural networks for pattern recognition and classification.[17] This application is promising because of its robust characteristics in dealing with noise and detecting damage-caused signals. It is capable of detecting patterns of features that define the class of unforeseen signals with interference. Accordingly, the objective of this study is to investigate the effectiveness of this approach for damage detection of composites. Being used as a nondestructive testing (NDT) technique, damage of detection composites is accomplished using wavelet neural networks with supervised training, and making use of decomposition of wavelet packets as a physical parameter in the identification. Another advantage over traditional NDT techniques is the possibility of identifying whether internal failure can also be made of plastics.

In Secs. II–IV we briefly review the orthonormal wavelet packets' decomposition and the autoregression (AR) spectrum estimation based on decomposing the wavelet packets. Section V presents the overview of the wavelet neural network, which is applied to classification. In Sec. VI we draw our strategy of data fusion and experimental design for the recognition of different damage of compound beams, i.e., no damage, cracks, and delaminations with different sizes. Finally, Sec. VII summarizes the results and conclusions of our study.

Orthonormal Wavelet Packets

Wavelet packets proposed by Coifman and Meyer[14] are a generalization of orthonormal wavelets. Consider a real-valued mother

Received 21 October 1998; revision received 28 September 1999; accepted for publication 7 October 1999.

*Associate Professor, Department of Biomedical Engineering.
†Professor, State Key Laboratory of Vibration Shock and Noise.
‡Professor, Department of Biomedical Engineering.

orthonormal wavelet function ψ and the associated scaling function ϕ. Let p_k define the filter coefficients of the scaling function ϕ, and $q_k = (-1)^k p_{1-k}$ correspond to that of the wavelet function ψ. We denoted

$$\mu_0(x) = \phi(x), \qquad \mu_1(x) = \psi(x) \tag{1}$$

$$P_0(z) = \frac{1}{2}\sum_k p_k z^k$$

$$P_1(z) = \frac{1}{2}\sum_k q_k z^k = \frac{1}{2}\sum_k (-1)^k \bar{p}_{-k+1} z^k \tag{2}$$

We now define the scaling function ϕ and the wavelet function ψ from the coefficients $\{p_k\}$ and $\{q_k\}$.

$$\mu_0(x) = \sum_k p_k \mu_0(2x-k), \qquad \mu_1(x) = \sum_k q_k \mu_0(2x-k) \tag{3}$$

Applying the Fourier transform to Eq. (3), the following relations are obtained:

$$\hat{\mu}_0(\omega) = P_0(e^{-i\omega/2})\hat{\mu}_0(\omega/2), \qquad \mu_1(\omega) = P_1(e^{-i\omega/2})\hat{\mu}_0(\omega/2) \tag{4}$$

where functions $\hat{\mu}_0(\omega)$ and $\hat{\mu}_1(\omega)$ is the Fourier transform of functions $\mu_0(x)$ and $\mu_1(x)$. In this new notation wavelet packets can be introduced more conveniently. These functions could show the local time-frequency properties.

Definition 1: Define the function μ_n, $n = 2l$ or $n = 2l+1$, $l = 0, 1, \ldots$

$$\mu_{2l}(x) = \sum_k p_k \mu_l(2x-k), \qquad \mu_{2l+1}(x) = \sum_k q_k \mu_l(2x-k) \tag{5}$$

then μ_n are described as wavelet packets that can be recursively calculated by the scaling function $\mu_0 = \phi$.

One can see[20] that such wavelet packets are orthonormal in $L^2(R)$ and they can serve as a set of bases similar to the orthonormal wavelet functions. Furthermore, wavelet packets are well localized in both time and frequency and thus provide an attractive alternative to time-frequency analysis.

Orthonormal Decompose of Wavelet Function

Suppose the wavelet packets can be shown by a group of functions $\{\mu_n\}$ obtained from an orthonormal scaling function $\mu_0 = \phi$. For each positive integer $n \in Z_+$, we define a group of subspace U_j^n obtained by $\{\mu_n\}$:

$$U_j^n = \mathrm{close}_{L^2(R)}\langle 2^{j/2}\mu_n(2^j \bullet - k)\rangle, \qquad k \in Z, j \in Z, n \in Z_+ \tag{6}$$

where the notation $\mathrm{close}_{L^2(R)}\langle\ \rangle$ indicates that U_j^n is the set of all possible finite linear combinations of $\mu_n(2^j \bullet - k)$ with fixed j, as well as their limits under the $L^2(R)$ norm, respectively, then the next relation can be concluded:

$$U_j^n = V_j, \qquad j \in Z$$

$$U_j^{n+1} = W_j, \qquad j \in Z \tag{7}$$

where the subspaces $\{V_j\} \in L^2(R)$ are generated by scaling function $\mu_0 = \phi$ and the orthogonal complementary subspaces $\{W_j\}$ of $\{V_j\}$ are generated by wavelet $\mu_1 = \psi$. The orthogonal decompositions are given[21]:

$$V_{j+1} = V_j \oplus W_j, \qquad j \in Z$$

where the sign $\oplus$ is a direct sum of the orthogonal decompositions of subspaces $\{V_j\}$ and $\{W_j\}$. The preceding equation is equivalent to

$$U_{j+1}^0 = U_j^0 \oplus U_j^1, \qquad j \in Z \tag{8}$$

We may inductively demonstrate that the preceding formulas are still correct for any positive integer $n \in Z_+$.

$$U_{j+1}^n = U_j^{2n} \oplus U_j^{2n+1}, \qquad j \in Z \tag{9}$$

An orthonormal base $\{\psi_{j,k}\}$ $j, k \in Z$ can be created by an orthonormal wavelet ψ, and for each $j \in Z$ the collection of the jth function $\{\psi_{j,k}\}$ $k \in Z$ is not only an orthonormal base in the subspace $W_j = \mathrm{close}_{L^2(R)}\langle \psi_{j,k}, k \in Z\rangle$, but also it can extract the local time-frequency information in the jth band:

$$H_j = \left(2^{j+1}\Delta_{\hat{\psi}}, 2^{j+2}\Delta_{\hat{\psi}}\right) \tag{10}$$

where $\Delta_{\hat{\psi}}$ is the bandwidth that corresponds to the mean quadratic root of wavelet. With respect to the higher frequency, the width H_j increases. We will show that wavelet packets have better frequency localization property and also better ability in dividing higher-frequency bandwidth.

Theorem 1: For each $j = 1, 2, \ldots$, the following equalities are held:

$$\begin{cases} W_j = U_{j-1}^2 \oplus U_{j-1}^3 \\ W_j = U_{j-2}^4 \oplus U_{j-2}^5 \oplus U_{j-2}^6 \oplus U_{j-2}^7 \\ \cdots\cdots \\ W_j = U_{j-k}^{2^k} \oplus U_{j-k}^{2^k+1} \oplus \cdots \oplus U_{j-k}^{2^{k+1}+1} \\ \cdots\cdots \\ W_j = U_0^{2^j} \oplus U_0^{2^j+1} \oplus \cdots \oplus U_0^{2^{j+1}+1} \end{cases} \tag{11}$$

Moreover, the library of functions

$$\left\{2^{(j-k)/2}\mu_{2^k+m}(2^{j-k}x - l), l \in Z\right\} \tag{12}$$

are a normal orthonormal basis in subspace $U_{j-k}^{2^k+m}$ for each $m = 0, \ldots, 2^k - 1$ $k = 1, \ldots, j$, and $j = 1, 2, \ldots$. In fact the kth orthonormal decomposition in the jth frequency band H_j can be divided into 2^k subbands

$$H_j^{k,m}, m = 0, \ldots, 2^k - 1 \tag{13}$$

The functions in Eq. (12) provide local time concentration in subbands $H_j^{k,m}$, and the combinations of set $\{H_j^{k,m}, m = 0, \ldots, 2^k - 1\}$ give the all-band H_j.

For any fixed value k $1 \le k \le j$, all wavelet series $g_j(x)$

$$g_j(x) = \sum_n d_n^j \psi(2^j x - n) = \sum_n d_n^j \mu_1(2^j x - n) \tag{14}$$

can be decomposed into an orthonormal sum of packets:

$$g_{j,k,m}(x) = \sum_n d_n^{j,k,m} \mu_{2^k+m}(2^{j-m}x - n), \qquad m = 0, \ldots, 2^k - 1 \tag{15}$$

The preceding formulas can be expressed as a binary tree, and the decomposition of the tree should be convenient in application. If the branch of some wavelet packets is less important than others, a small value k should be chosen for this branch.

When the value k is the largest in Eq. (11), we have the following:

Theorem 2: For each $j = 0, 1, 2, \ldots$, we have the decomposition

$$L^2(R) = \bigoplus_{j \in Z} W_j = \cdots \oplus W_{-1} \oplus W_0 \oplus U_0^2 \oplus U_0^3 \oplus \cdots \tag{16}$$

where the sets $\{\psi_{j,k}, \mu_n(\bullet - k)$, $j = \ldots, -1, 0; n = 2, 3, \ldots; k \in Z\}$ are normal orthonormal bases.

This process can be depicted schematically in Fig. 1. The operator H and G is constructed by the filter coefficients p_k of scaling function ϕ and the q_k of the wavelet function ψ, respectively.

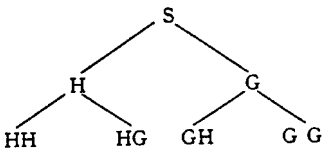

Fig. 1 Binary decomposition of wavelet packets.

The advantage of packets algorithm lies in the adaptive selection of the best bases according to the actual signals. Decomposition of wavelet packets is achieved by the algorithm of Quadrature Millor Filter, which projects signals to a set of orthogonal conjugate filters constructed by high-pass and low-pass filters. This approach can extract the essential features so that the wavelet neural network can distinguish different signals effectively. Based on the entropy function, we decompose the signals according to the sixth Daubechies compactly supported wavelet.[22] When the process is applied once, the data points keep one out of two, and the sampling interval doubles, as shown in Fig. 1. If the initial data contain 1024 points, the number of data points becomes fewer and fewer as the decomposition continues. As a consequence, we cannot correctly make out the frequency component by fast Fourier transform (FFT). In the next section an autoregressive algorithm will be used to solve the problem of short data sequence, which performs better than FFT analysis.

Foundation of AR Model and Spectrum Analysis

Suppose that we have the data $\{x_k\}\, k = 1, 2, \ldots, N$ decomposed by wavelet packets, we may then use the AR model to perform the spectrum analysis. The order p of the autoregressive model is determined by experiment or Akaike's information criteria.[23] The Yule–Walker equation that describes the relations of AR signal model parameters and autocorrelation can be easily derived. If the order p of AR is assumed, we have

$$x_n = -\sum_{l=1}^{p} a_l x_{n-l} + w_n \tag{17}$$

where w_n is white noise with zero mean and covariance σ^2. From Eq. (17) we can get

$$\begin{aligned} r_{xx}(k) &= E(x_{n+k}\overline{x_n}) \\ &= E\left[\left(-\sum_{l=1}^{p} a_l x_{n+k-l} + w_{n+k}\right)\overline{x_n}\right] \\ &= -\sum_{l=1}^{p} a_l r_{xx}(k-l) + E(w_{n+k}\overline{x_n}) \\ &= \begin{cases} -\sum_{l=1}^{p} a_l r_{xx}(k-l), & k > 0 \\ -\sum_{l=1}^{p} a_l r_{xx}(k-l) + \sigma^2, & k = 0 \end{cases} \end{aligned} \tag{18}$$

This is the Yule–Walker equation, and its parameter will be obtained next. First the p parameters $\{a_1, a_2, \ldots, a_p\}$ can be solved from the part $k > 0$ in Eq. (18). Then σ^2 is obtained from the part corresponding to $k = 0$. When the p autocorrelation functions are known, it is obvious that the solutions can be obtained. The Yule–Walker equation can be expressed as the following matrix form:

$$\begin{pmatrix} r_{xx}(0) & r_{xx}(-1) & \cdots & r_{xx}(-p+1) \\ r_{xx}(1) & r_{xx}(0) & \cdots & r_{xx}(-p+2) \\ \vdots & \vdots & & \vdots \\ r_{xx}(p-1) & r_{xx}(p-2) & \cdots & r_{xx}(0) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix} = -\begin{pmatrix} r_{xx}(1) \\ r_{xx}(2) \\ \vdots \\ r_{xx}(p) \end{pmatrix} \tag{19}$$

According to the property of autocorrelation functions, $r(-k) = \overline{r(k)}$. The autocorrelation matrix in Eq. (19) can be defined as

$$R_{xx} = \begin{pmatrix} r_{xx}(0) & r_{xx}(-1) & \cdots & r_{xx}(-p+1) \\ r_{xx}(1) & r_{xx}(0) & \cdots & r_{xx}(-p+2) \\ \vdots & \vdots & & \vdots \\ r_{xx}(p-1) & r_{xx}(p-2) & \cdots & r_{xx}(0) \end{pmatrix}$$

where matrix R_{xx} is a Hermitian satisfied $R_{xx}^H = R_{xx}$. Moreover it is also a Toeplitz matrix because of its diagonal elements being the same. If X_n is not pure harmonic quantity, the matrix R_{xx} has solutions $\{a_1, a_2, \ldots, a_p\}$. We add Eq. (18) (with $k = 0$) and Eq. (19) together and get

$$\begin{pmatrix} r_{xx}(0) & r_{xx}(-1) & \cdots & r_{xx}(-p) \\ r_{xx}(1) & r_{xx}(0) & \cdots & r_{xx}(-p+1) \\ \vdots & \vdots & & \vdots \\ r_{xx}(p) & r_{xx}(p-1) & \cdots & r_{xx}(0) \end{pmatrix} \begin{pmatrix} 1 \\ a_1 \\ \vdots \\ a_p \end{pmatrix} = -\begin{pmatrix} \sigma^2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{20}$$

To determine the variables $\{a_1, a_2, \ldots, a_p\}$ and σ^2, it is necessary to solve Eq. (20) by calculating $p + 1$ autocorrelations $r_{xx}(0), r_{xx}(1), \ldots, r_{xx}(p)$. If the variables of AR are obtained, the AR spectrum of decomposition of wavelet packets in the ith sequence of the lth layer is expressed as follows:

$$P_{\mathrm{AR}}(f - f_{\mathrm{li}}) = \frac{\sigma^2 / f_{\mathrm{ls}}}{\left|1 - \sum_{k=1}^{p} a_k \{\exp[-j2\pi(f - f_{\mathrm{li}})/f_{\mathrm{ls}}]\}\right|^2} \tag{21}$$

where the sampling frequency f_{ls} and initiative frequency f_{li} are related to the layer number of the wavelet decomposition. If the sampling frequency of the initial data is f_{Ns}, the data length is N, and we get 2^l sequences of different bands in the lth decomposition layer. The corresponding sampling frequency is $f_{\mathrm{ls}} = 2^{-l} f_{Ns}$, and the length of each sequence is $N_l = 2^{-l} N$. The initial frequency is $f_{\mathrm{li}} = f_{\mathrm{ls}} * (i - 1)/2$ by the theory of wavelet packets.

Adaptive B-Spline Wavelet Network

A wavelet network $g : R^m \rightarrow R^n$ can be expressed in the form[17]

$$\hat{g}_i(t) = \sigma\left\{\sum_{j=1}^{j_0} w_{ij}\left[\sum_{k=1}^{m} c_k(t)\psi\left(\frac{k - b_j}{a_j}\right)\right]\right\}, \qquad i = 1, 2, \ldots, n \tag{22}$$

where $c_k(t)(k = 1, 2, \ldots, m)$ is the kth input training sample, meaning the extracted feature of the signal according to the preceding theory. $\hat{g}_i$ $(i = 1, 2, \ldots, n)$ represents the ith output, and j_0 is the number of the hidden layers. The B-spline wavelet functions $\psi[(k - b_j)/a_j]$ serve to be the connective function from the kth input to the jth hidden layer, a_j and b_j are dilations and shifts for each wavelet, w_{ij} are the weight coefficients from the jth hidden layer to the ith output, and $\sigma(z)$ is a sigmoidal activative function. The expression (22) can be realized by the network shown in Fig. 2. The network parameters a_j, b_j, and w_{ij} can be optimized by minimizing an energy function according to the input samples $c_k(t)$ and training samples $g_j(t)$. The least-mean-squares criterion is employed as

$$E = \frac{1}{2}\sum_{p=1}^{p_0}\sum_{i=1}^{n}\left[g_i^p(t) - \hat{g}_i^p(t)\right]^2 \tag{23}$$

where p_0 is the total number of samples. By introducing momentum factor α for speeding up the convergence and stabilizing the change of the weights, the formulas of the change of the weights are given by

$$a_j(n_0 + 1) = a_j(n_0) - \eta\frac{\partial E}{\partial a_j} + \alpha\Delta a_j(n_0) \tag{24}$$

$$b_j(n_0 + 1) = b_j(n_0) - \eta\frac{\partial E}{\partial b_j} + \alpha\Delta b_j(n_0) \tag{25}$$

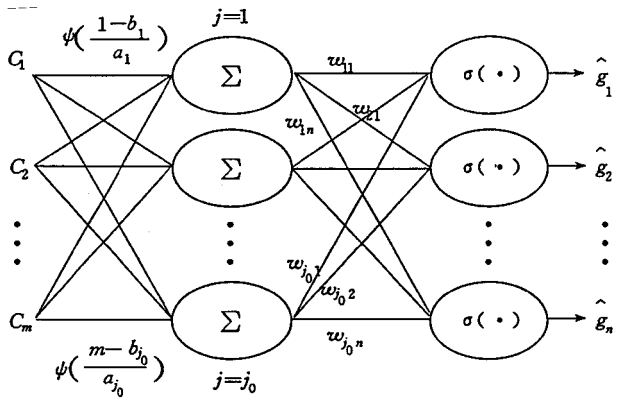

Fig. 2 Adaptive B-spline wavelet network.

$$w_{ij}(n_0 + 1) = w_{ij}(n_0) - \eta\frac{\partial E}{\partial w_{ij}} + \alpha\Delta w_{ij}(n_0) \tag{26}$$

where η is a variable step size and

$$\frac{\partial E}{\partial a_j} = \sum_{p-1}^{p_0}\sum_{i=1}^{n}\sum_{k=1}^{m}\left[g_i^p(t) - \hat{g}_i^p(t)\right] \times \sigma'\left\{\sum_{j=1}^{j_0} w_{ij}\left[\sum_{k=1}^{m} c_k^p(t)\psi\left(\frac{k-b_j}{a_j}\right)\right]\right\} w_{ij} c_k^p(t) \times \psi'\left(\frac{k-b_j}{a_j}\right)(k-b_j)\frac{1}{a_j^2} \tag{27}$$

$$\frac{\partial E}{\partial b_j} = \sum_{p-1}^{p_0}\sum_{i=1}^{n}\sum_{k=1}^{m}\left[g_i^p(t) - \hat{g}_i^p(t)\right] \times \sigma'\left\{\sum_{j=1}^{j_0} w_{ij}\left[\sum_{k=1}^{m} c_k^p(t)\psi\left(\frac{k-b_j}{a_j}\right)\right]\right\} w_{ij} c_k^p(t) \times \psi'\left(\frac{k-b_j}{a_j}\right)\frac{-1}{a_j} \tag{28}$$

$$\frac{\partial E}{\partial w_{ij}} = \sum_{p-1}^{p_0}\sum_{i=1}^{n}\sum_{k=1}^{m}\left[g_i^p(t) - \hat{g}_i^p(t)\right] \times \sigma'\left\{\sum_{j=1}^{j_0} w_{ij}\left[\sum_{k=1}^{m} c_k^p(t)\psi\left(\frac{k-b_j}{a_j}\right)\right]\right\} c_k^p(t)\psi\left(\frac{k-b_j}{a_j}\right) \tag{29}$$

Those are obtained from Eqs. (22) and (23).

Experimental Results

The proposed methodology will be presented to approach the practical damage detection of composite materials through the study of a simple example. We have made six composite beams with identical size using the same materials and technologies. They are of the same length, thickness and width, 60.0 ×16.0 ×2.5 cm, respectively. The first beam contains no crack and delamination, and it is classified as health. Three of the beams contain crack of sizes 2.0, 2.5, and 3.0 cm; two of the beams have delaminations of 1.5 ×1.5 cm and 2.0 ×2.0 cm. All of the cracks and delaminations are located in the middle of the beams.

Figure 3 shows the experimental setup. A piezoelectric ceramic patch, functioning as an actuator, was bonded to the top of each beam 6.0 cm from a clamp using super glue. A second piezoelectric ceramic patch, functioning as a sensor, was bonded in the same manner, 6.0 cm from the opposite end of the beam. Each beam is clamped 2.5 cm from the end closest to the actuating piezoelectric ceramic patch. The cantilever beam length is 60.0 cm. The natural

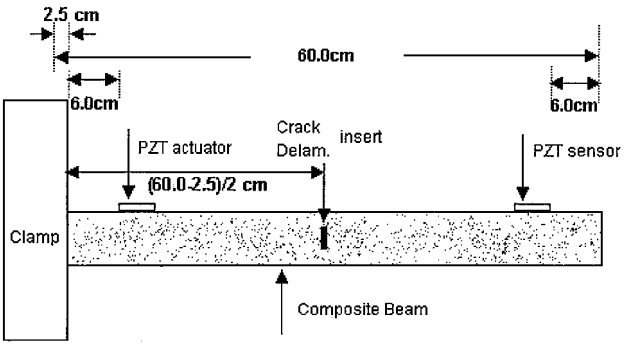

Fig. 3 Experimental setup for detection of a cantilever composite beam with prescribed damage.

frequencies of health specimen are found using piezoelectric ceramic sensors by an experiment. The first two modal frequencies are too low, so we choose the third frequency 1.0 kHz to stimulate the cantilever beam. We find that the cantilever beams have a more stable state of vibration with a narrow bandwidth using the 1.0-kHz square wave as the center frequency than using that of the 1.0-kHz sine wave. The square-wave signal drives the actuating piezoelectric ceramic patch after amplification by Piezo Systems. The output acceleration signal of the health compounds, measuring by the piezoelectric ceramic patch sensor, is sampled and shown in Fig. 4a, whereas those of the other classes are plotted in Figs. 5a–9a, respectively. The corresponding spectrum of Fourier transform are shown in Figs. 4b–9b. Moreover, the corresponding eight figures of decomposition of packets in the third layer, obtained by decomposition using the sixth Daubechies[22] compactly supported wavelet, are shown in Figs. 4c–9c. The corresponding eight frequency spectra plots obtained from the decomposition of packets in the third layer are shown in Figs. 4d–9d. The available information is not fully consistent for cracks with almost the same degree of damage because of complexity in making the structure of compounds, dispersions, and random disturbance in the process of measurement. Signal of interest can be obscured by background signal sources. From the Fourier transform spectra shown in Figs. 4b–9b, the health material is found to have some main peaks, whereas those of the cracks and deliminations samples also have some. All of these peaks lie in different frequencies. These characteristics might be used to identify the composites from damaged ones. But the frequencies of main peaks lie in different frequencies for cracks and deliminations with different sizes in Figs. 5b–9b. The structure of composites with dispersion may be unstable, inconsistent, and complex in behavior. As a result, it is quite difficult to obtain a set of consistent features. To improve the fault tolerance and interference immunity, we should take into account the transmission bands, extract, and distinguish the features in different frequency bands before we can obtain the essential feature corresponding to each damage.

Because of the unstationary property, it is better to manipulate these signals using wavelet packets than Fourier transform. Figures 4d–9d show the eight AR spectra of the wavelet packets corresponding to each beam. From the figures we observe that the mean values of energy spectrum of signals in different bands are almost the same for the same type damage, although there are deviations in position of the peaks. However, they vary largely for different types of damage. This mean value of energy spectrum reflects the features of damage types and shows the inner property attributed by different damages. We extract these mean values as the features of the signals from Figs. 4d–9d. Eight feature data $c_k(t)$ $(k = 1, 2, \ldots, 8)$ are extracted in every case and are listed as Table 1. They can then be input to the adaptive wavelet network for training and classifying. The network consists of eight input features, eight hidden nodes, and three output nodes corresponding to the three classes of the damage composite, i.e., health, cracks and delaminations of different degrees. When the features $c_k(t)$ $(k = 1, 2, \ldots, 8)$ corresponding to the health class are fed to the network, the output $\hat{g}_i$ $(i = 1, 2, 3)$ are defined as (1,0,0). When those correspond to the crack class and delaminations are fed, the output $\hat{g}_i$ $(i = 1, 2, 3)$ are respectively defined as (0,1,0) and (0,0,1). In the network ψ is the fourth B-spline

Table 1 Eight mean values of decomposition of energy spectrum

Type	First AR mean	Second AR mean	Third AR mean	Fourth AR mean	Fifth AR mean	Sixth AR mean	Seventh AR mean	Eighth AR mean
Health	1.0913	1.0808	0.1744	0.5127	0.0548	0.1763	0.0494	0.1060
2.0-cm crack	0.7028	0.4507	0.1613	0.3566	0.0535	0.1353	0.0449	0.0870
2.5-cm crack	0.6691	0.3583	0.1150	0.3515	0.0407	0.1347	0.0519	0.0782
3.0-cm crack	0.6766	0.4149	0.0937	0.1633	0.0339	0.1089	0.0508	0.0709
1.5 ×1.5 cm delam.	0.1967	0.1336	0.0786	0.0654	0.0138	0.0277	0.0338	0.0255
2.0 ×2.0 cm delam.	0.1102	0.0827	0.0720	0.0372	0.0126	0.0238	0.0312	0.0213

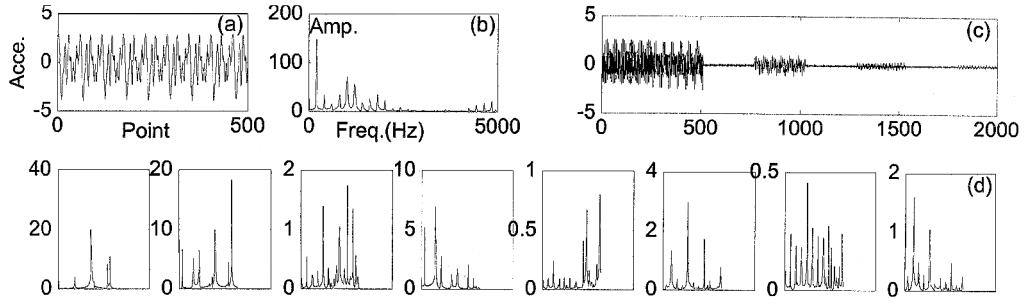

Fig. 4 Health composite: a) output signal, b) amplitude-frequency of the output sampling signal, c) eight decomposition of packets in third layer, and d) corresponding eight AR spectrum plots according to the decomposition of packets in the third layer.

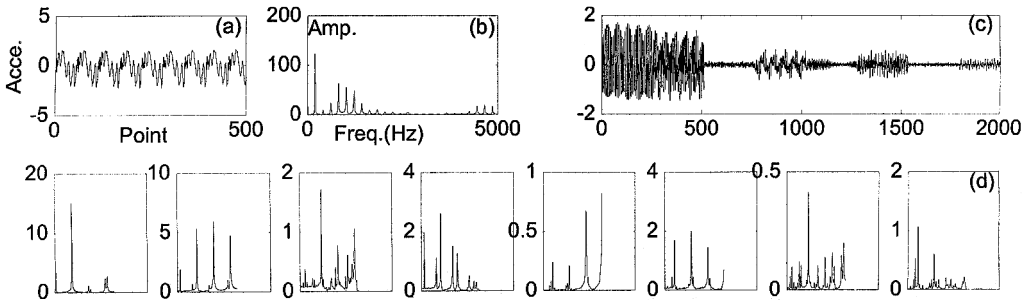

Fig. 5 Crack size of 2.0 cm: a) output sampling signal, b) amplitude frequency of the output sampling signal, c) eight decomposition of packets in third layer, and d) corresponding eight AR spectrum plots according to the decomposition of packets in third layer.

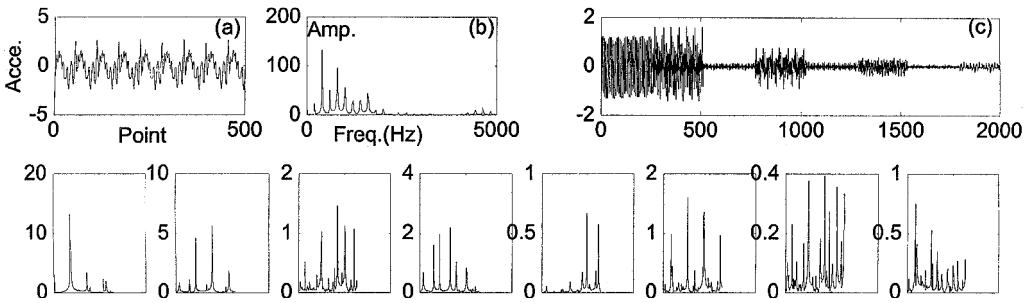

Fig. 6 Crack size of 2.5 cm: a) output sampling signal, b) amplitude frequency of the output sampling signal, c) eight figures of decomposition of packets in third layer, and d) corresponding eight AR spectrum plots according to the decomposition of packets in third layer.

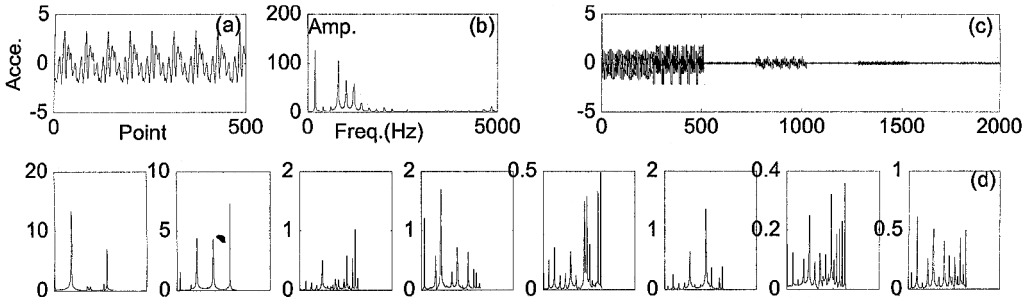

Fig. 7 Crack size of 3.0 cm: a) output sampling signal, b) amplitude frequency of the output sampling signal, c) eight decomposition of packets in third layer, and d) corresponding eight AR spectrum plots according to the decomposition of packets in third layer.

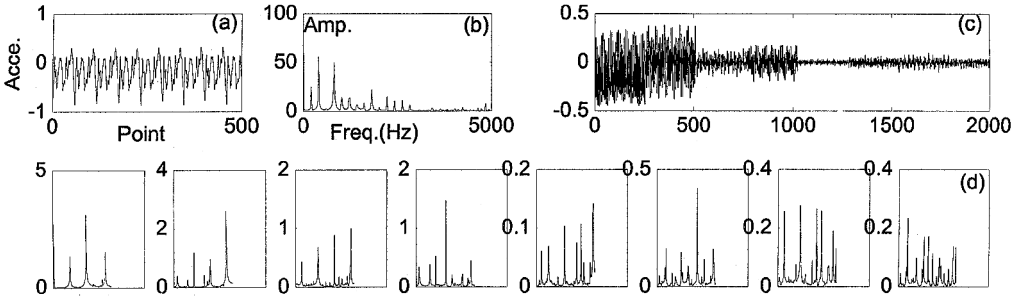

Fig. 8 Delamination size of 1.5 × 1.5 cm: a) output sampling signal, b) amplitude frequency of the output sampling signal, c) eight decomposition of packets in third layer, and d) corresponding eight AR spectrum plots according to the decomposition of packets in third layer.

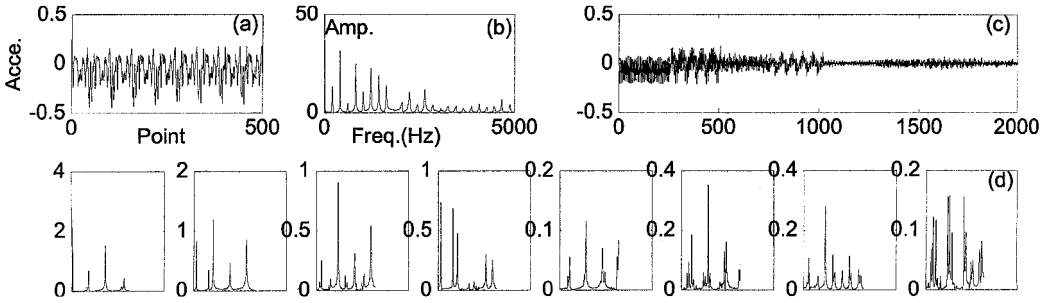

Fig. 9 Delamination size of 2.0 × 2.0 cm: a) output sampling signal, b) amplitude frequency of the output sampling signal, c) eight decomposition of packets in third layer, and d) corresponding eight AR spectrum plots according to the decomposition of packets in third layer.

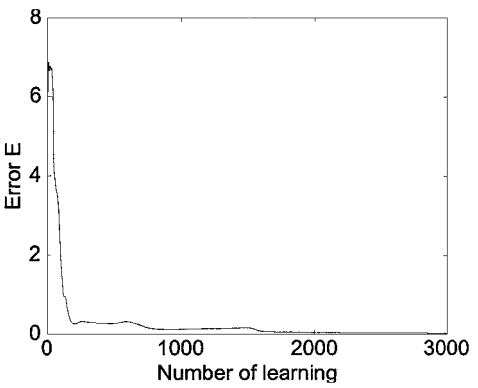

Fig. 10 Learning convergent curve of the wavelet neural network.

wavelet[19] empirically. Because the neural network is simple, the error is less than 10^{-2} after 3000 batch-mode iterations of the gradient descent algorithm using Eqs. (24–26). The learning convergent curve of the network is shown in Fig. 10. After training the other composites made corresponding to the three states are tested. The features of the signals extracted by the wavelet packets are fed to wavelet neural network. Generally, only one of the output patterns $\hat{g}_i$ $(i = 1, 2, 3)$ is equal to 1. If there are more outputs equal to 1, we choose the output pattern with the maximal argument. The performance of the network that classifies the signals corresponds to different classes is correct. Moreover the system can also recognize signals with small perturbance.

Conclusion

Testing signals from composites with different types of damage have different behaviors in different bands. Wavelet packets project a signal into a set of basic wavelet functions so that the signals can be exactly decomposed to different frequency bands. We extract the useful information in different bands for distinguishing different levels of damage. Further research will be done on a series of test samples using wavelet packets to study composites to identify the location and magnitude of the damage in addition to its type.

Acknowledgments

This research work was financially supported by a research grant from National Natural Science Fund of China and Aeronautics Science Fund. We are grateful to BaoQi Tao for his help in performing the experiment.

References

[1]Jaewook, R., and Lee, S. W., "A Network Approach for Damage Detection and Identification of Structures," AIAA Paper 94-1753, April 1994.

[2]Manning, R. A., "Damage Detection in Adaptive Structures Using Neural Networks," AIAA Paper 94-1752, April 1994.

[3]Ni, Y. Q., Wang, B. S., and Ko, J. M., "Selection of Input Vectors to Neural Networks for Structural Damage Identification," *Proceedings of SPIE—The International Society for Optical Engineering*, Vol. 3671, Society of Photo-Optical Instrumentation Engineers, Bellingham, WA, 1999, pp. 270–280.

[4]Cios, K. J., and Tjia, R. E., "Application of Neural Networks in the Acousto-Ultrasonic Evaluation of Metal-Matrix Composite Specimens," *International Joint Conference on Neural Networks*, International Neural Network Society and Inst. of Electrical and Electronics Engineers, Singapore, Vol. 2, 1992, pp. 993–998.

[5]Mallat, S. G., "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, 1989, pp. 674–693.

[6]Daubechies, I., "The Wavelet Transform, Time Frequency Localization and Signal Analysis," *IEEE Transactions on Information Theory*, Vol. 36, No. 5, 1990, pp. 961–1005.

[7]Dalpiaz, G., and Rivola, A., "Condition Monitoring and Diagnostics in Automatic Machines: Comparison of Vibration Analysis Technique," *Mechanical Systems and Signal Processing*, Vol. 11, No. 1, 1997, pp. 53–73.

[8]Staszewski, W. J., and Worden, K., "Classification of Faults in Gearboxes—Pre-Processing Algorithms and Neural Networks," *Neural Computing and Applications*, Vol. 5, No. 3, 1997, pp. 160–183.

[9]Yang, J., Xiang, X., and Xiong, S., "Damage Detection of Roller Bearing Using Wavelet Transform," *Proceedings of the International Modal Analysis Conference*, Vol. 2, Society for Experimental Mechanics, 1999, pp. 1093–2085.

[10]Staszewski, W. J., "Structural and Mechanical Damage Detection Using Wavelets," *Shock and Vibration Digest*, Vol. 30, No. 6, 1998, pp. 457–472.

[11]Deng, X. M., and Wang, Q., "Crack Detection Using Spatial Measurements and Wavelet Analysis," *International Journal of Fracture*, Vol. 91, No. 2, 1998, pp. L23–L28.

[12]Lu, C. J., and Hsu, Y. T., "Application of Wavelet Transform to Structural Damage Detection," *Proceedings of the International Modal Analysis*

Conference, Vol. 1, Society for Experimental Mechanics, 1999, pp. 908–914.

[13]Steinberg, B. Z., McCoy, J. J., and Mirotznik, M., "A Wavelet Approach to Effective Modal Analysis of Complex Lamination," *IEEE International Symposium Antennas and Propagation Society*, Vol. 2, Inst. of Electrical and Electronics Engineers, Piscataway, NJ, 1997, pp. 1096–1099.

[14]Coifman, R. R., and Meyer, Y., "Orthonormal Wavelet Packet Bases," Dept. of Electrical Engineering, TR, Yale Univ., New Haven, CT, Aug. 1989.

[15]Coifman, R. R., and Wickerhauser, M. V., "Entropy-Based Algorithms for Best Basis Selection," *IEEE Transactions on Information Theory*, Vol. 38, No. 5, 1992, pp. 713–718.

[16]Herley, C., Kovacelevic, J., Ramchandran, K., and Vetterli, M., "Tilings of the Time-Frequency Plane: Construction of Arbitrary Orthogonal Bases and Fast Tiling Algorithms," *IEEE Transaction Signal Processing*, Vol. 41, No. 12, 1993, pp. 3341–3359.

[17]Szu, H. H., Telfer, B., and Kadambe, S., "Neural Network Adaptive Wavelets for Signal Representation and Classification," *Optical Engineering*, Vol. 31, No. 9, 1992, pp. 1907–1916.

[18]Wong, K. W., and Leung, C. S., "On-Line Successive Synthesis of Wavelet Networks," *Neural Processing Letters*, Vol. 7, No. 3, 1998, pp. 91–100.

[19]Wu, Y., Tao, B., and Yan, S., "B-Spline Wavelet Neural Network," *Pattern Recognition and Artificial Intelligence*, Vol. 9, No. 3, 1996, pp. 228–233 (in Chinese).

[20]Chui, C. K., and Wang, J. Z., "On Compactly Supported Spline Wavelets and a Duality Principle," *Transactions of the American Mathematical Society*, Vol. 330, No. 2, 1992, pp. 903–915.

[21]Chui, C. K., *An Introduction to Wavelets*, Academic, Boston, 1992, pp. 236–243.

[22]Daubechies, I., "Orthonormal Bases of Compactly Supported Wavelets," *Communication in Pure and Applied Mathematics*, Vol. 41, No. 8, 1988, pp. 520–564.

[23]Akaike, H., "Power Spectrum Estimation Through Autoregression Model Fitting," *Annals of the Institute of Statistical Mathematics*, Vol. 21, 1969, pp. 407–419.

G. A. Kardomateas
Associate Editor