Available online at www.sciencedirect.com
ScienceDirect
Journal of Sound and Vibration 321 (2009) 590–604
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

Structural damage localization using wavelet-based silhouette statistics

Uk Jung[a], Bong-Hwan Koh[b,*]

[a]Department of Business Administration, Dongguk University, 3-26 Pil-dong, Chung-gu, Seoul 100-715, Republic of Korea
[b]Department of Mechanical Engineering, Dongguk University, 3-26 Pil-dong, Chung-gu, Seoul 100-715, Republic of Korea

Received 3 September 2007; received in revised form 6 March 2008; accepted 19 October 2008
Handling Editor: L.G. Tham
Available online 29 November 2008

Abstract

This paper introduces a new methodology for classifying and localizing structural damage in a truss structure. The application of wavelet analysis along with signal classification techniques in engineering problems allows us to discover novel characteristics that can be used for the diagnosis and classification of structural defects. This study exploits the data discriminating capability of silhouette statistics, which is eventually combined with the wavelet-based vertical energy threshold technique for the purpose of extracting damage-sensitive features and clustering signals of the same class. This threshold technique allows us to first obtain a suitable subset of the extracted or modified features of our data, i.e. good predictor sets should contain features that are strongly correlated to the characteristics of the data without considering the classification method used, although each of these features should be as uncorrelated with each other as possible. The silhouette statistics have been used to assess the quality of clustering by measuring how well an object is assigned to its corresponding cluster. We use this concept for the discriminant power function used in this paper. The simulation results of damage detection in a truss structure show that the approach proposed in this study can be successfully applied for locating both open- and breathing-type damage even in the presence of a considerable amount of process and measurement noise. Finally, a typical data mining tool such as classification and regression tree (CART) quantitatively evaluates the performance of the damage localization results in terms of the misclassification error.

1. Introduction

Evaluating the structural integrity of engineering structures has been a critical research topic for the last few decades due to the ever-increasing demands for longer service life. Most engineering structures are easily exposed to unforeseen defects internally and externally such as fatigue crack, thermal degradation, impact, overloading, and corrosion under normal operating conditions. Since the overall cost of repair or downtime can be significant for severely damaged social infrastructure, the importance of early detection and damage forecasting has increased considerably. The oldest and simplest measure of damage detection, i.e. visual inspection, is feasible only for examining human-accessible areas, which is impossible under certain

*Corresponding author.
E-mail address: bkoh@dongguk.edu (B.-H. Koh).

0022-460X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2008.10.016

circumstances. Therefore, structural health monitoring systems that use measurements based on the vibration characteristics of a structure have attracted attention in recent years [1,2]. Many vibration-based damage detection methods exploit the changes that occur in dynamic properties such as modal parameters (natural frequencies, mode shapes, and frequency response functions) due to stiffness or mass variations in a structure, so that the structural integrity can be remotely and indirectly assessed without requiring human inspection. However, one of the drawbacks of vibration-based approaches is the inherently low sensitivity of the modal properties toward small-sized structural defects [3]. It is very difficult to determine the variations in damage-causing modal properties until the severity of damage becomes significant. Some research efforts have attempted to address this low-sensitivity issue by employing feedback control in a smart structure framework [4]. However, implementing active control systems in a host structure remains an expensive option in most cases.

Recently, emerging technologies in data processing have raised the possibility of using the concept of wavelet transformation in solving structural health monitoring problems [5–7]. Many previous vibration-based methods heavily relied on the use of Fourier transform, which decomposes a signal into constituent sinusoids of different frequencies. The Fourier transform basically extracts frequency components from a time domain signal. Consequently, during the transformation, the time information is lost and it is impossible to determine when or where a particular event occurred. Although various types of windowing techniques such as short-time Fourier transform can be used to overcome this problem, the general level of precision is also limited by the size of the employed window. An alternative measure is to use wavelet transforms. The main advantage of wavelet transform is its ability to perform fine-scale analysis on highly localized signals such as spikes or abnormalities. In this respect, wavelets can be contained in a finite interval. Hence, they are suitable for representing and approximating a signal having many abrupt variations. These advantages are considered to be desirable features for structural damage detection applications, particularly when the Fourier transform suffers from low sensitivity because of its inherent ''averaging-out'' effect. Moreover, due to the availability of fast transform algorithms, the computational cost of performing the signal transformation has been reduced significantly. Because of these features, wavelet transform is considered to be one of the most powerful tools for structural health monitoring.

Moyo and Brownjohn proposed a wavelet strain model for bridge condition monitoring, where a rapid drop in strain readings can be considered as a potential cause of cracks in reinforced concrete structures [8]. Their study makes use of the wavelet transform of strain data that is recorded as early as the construction stage, and attempts to decompose the strain time-history into multiple time-frequency components. This transform allows the separation of damage-induced features from other harmless influences. Hou et al. [5] applied wavelet transform to detect abrupt changes in the stiffness in a simple spring–mass–damper system. Their study showed that the process of accumulative structural degradation can be identified as spikes in level 1 details of wavelet decomposition. Sun and Chang [9] proposed a combined method using wavelet packet transforms and neural networks for damage assessment in a three-span bridge model. The wavelet packet sensitivity is also used for detecting damage in a simply supported concrete beam. In other words, this study exploits the correlation between the wavelet packet sensitivity and the local change in system parameters. Moreover, Li et al. [10] first decomposed the structural response data into several empirical modes. These monocomponent signals were further processed through wavelet transforms to identify the stiffness damage in a four-storey shear building model. Lu and Hsu [11] investigated a wavelet-based approach for detecting a nonuniformly added mass on a string system. Basu [12] showed that a wavelet-based approach allows the detection of the bilinear behaviour of nonhysteretic systems. Staszewski [7] summarized some of the developments and applications in wavelet-based damage detection research. Jung et al. [13] developed a vertical-energy-based thresholding (VET) procedure. This wavelet feature selection procedure balances the reconstruction error against the data-reduction efficiency. It is effective in capturing the key patterns in multiple data signals and removing noise such that the selected wavelet coefficients are treated as the reduced number of features in subsequent analysis for decision making such as clustering and classification.

This paper investigates a structural damage localization problem using a wavelet-based signal classification method that extracts signal features with the best discrimination ability when classifying the location of stiffness damage in a planar truss structure. The proposed approach uses simulation data generated from a truss model subjected to an unknown random excitation. Since most signal features in the damage-induced

response are irrelevant to the class distinction and inevitably corrupted with measurement noises, we first attempt to apply the VET criteria previously proposed by Jung et al. [13]. According to these criteria, good predictor sets should contain features that are strongly correlated to the class distinction, although each of these features should be as uncorrelated with each other as possible; we thus select differentiated features for data dimensionality reduction and noise removal. Secondly, the proposed VETS (VET wavelet positions containing large silhouette statistics) comprises a few features with highest silhouette statistics to find the smaller number of features having more discriminating power for localizing a stiffness-damaged element in a truss structure. Finally, the VETS procedure is applied to a commonly used data mining tool (classification and regression tree (CART) [14]) for achieving signal classification and obtaining the misclassification error to quantify the overall performance of damage localization.

2. Review of related theories

2.1. Wavelet transformation

This section briefly reviews the theoretical background of discrete wavelet transform (DWT). See Mallat [15] for more details. DWT effectively projects a temporal signal into a special wavelet basis that entails adjustable multiresolution parameters such as scale and position to represent a nonstationary signal. Typically, DWT is performed on multiple levels with different frequency resolutions. As each level of the transformation is performed, there is a trade-off between the time and frequency resolution. The full DWT for a time domain signal in L_2 (finite energy), $f(t)$, can be represented in terms of a shifted version of a scaling function $\phi(t)$ and a shifted and dilated version of a so-called mother wavelet function $\psi(t)$. DWT can be represented as

$$f(t)=\sum_{k\in Z} c_{L,k}\phi_{L,k}(t)+\sum_{j\geqslant L}\sum_{k\in Z} d_{j,k}\psi_{j,k}(t) \tag{1}$$

where $d_{j,k}$ are the wavelet coefficients and $c_{L,k}$, $L<J$ are the scaling coefficients. These coefficients are given by the inner product in L_2, i.e.

$$c_{L,k}=\langle f(t),\phi_{L,k}(t)\rangle \quad \text{and} \quad d_{j,k}=\langle f(t),\psi_{j,k}\rangle \tag{2}$$

Here, $\phi_{L,k}(t)=2^{L/2}\phi(2^L t-k)$; $k\in Z$ is a family of scalar functions and $\psi_{j,k}(t)=2^{j/2}\psi(2^j t-k)$; $j\geqslant L$, $k\in Z$ is a family of wavelet functions. If the mother functions are properly selected, their family forms an orthogonal basis for the signal space.

Consider a sequence of data $\mathbf{y}=(y(t_1),\ldots,y(t_N))$ taken from $f(t)$ or obtained as a realization of $y(t)=f(t)+\varepsilon_t$ at equally spaced discrete time points $t=t_i s$, where $\varepsilon_{t_i} s$ are independent and identically distributed (i.i.d.) noises following $N(0,\sigma^2)$. The DWT of $\mathbf{y}$ is defined as $\mathbf{d}=W\mathbf{y}$, where W is the orthonormal $N\times N$ DWT matrix. It is given that $\mathbf{d}=(\mathbf{c}_L,\mathbf{d}_L,\mathbf{d}_{L+1},\ldots,\mathbf{d}_J)$, where $\mathbf{c}_L=(c_{L,0},\ldots,c_{L,2^L-1})$, $\mathbf{d}_L=(d_{L,0},\ldots,d_{L,2^L-1})$, and $\mathbf{d}_J=(d_{J,0},\ldots,d_{J,2^J-1})$. Using inverse DWT, the $N\times 1$ vector $\mathbf{y}$ of the original signal curve can be reconstructed as $\mathbf{y}W^{\mathrm{T}}\mathbf{d}$. By applying DWT to the data $\mathbf{y}s$, $\mathbf{d}=W\mathbf{y}$, we obtain the following model in the wavelet domain: $d_{j,k}=\theta_{j,k}+\eta_{j,k}$ for $j=L,\ldots,J$, $k=0,1,\ldots,2^j-1$, and $c_{L,k}=\theta_{L,k}+\eta_{L,k}$ for $k=0,1,\ldots,2^L-1$, where $J=\log_2 N-1$. The model can be represented in the vector format as follows:

$$\mathbf{d}=\theta+\eta \tag{3}$$

where $\mathbf{d}$, θ, and η represent the collection of all coefficients, parameters, and errors, respectively. Since W is an orthonormal transform, $\eta_{j,k}$'s are still i.i.d. $N(0,\sigma^2)$ [16]. In order to simplify the notation used in this paper, $\mathbf{d}=(d_1,d_2,\ldots,d_N)$ is used instead of c_{Lk} and d_{jk} for the components of $\mathbf{d}$.

2.2. Wavelet model for multiple signals

We denote a vector of N equally-spaced data points from a signal curve, where $N=2^J$ with some positive integer J and $i=1,2,\ldots,M$ by $y_i=[y_{i1},y_{i2},\ldots,y_{iN}]$. Let $Y=[y_1,y_2,\ldots,y_M]^{\mathrm{T}}$ be the collection of M multiple sets of functional data. When DWT W is applied to a data set, the matrix of wavelet coefficients obtained from

this transformation is $D = WY$, where $D = [d_1, d_2, \ldots, d_M]^{\mathrm{T}}$, $d_i = [d_{i1}, d_{i2}, \ldots, d_{iN}]$, and d_{im} is the wavelet coefficient at the mth wavelet position for the ith data curve. The model of wavelet coefficients D from M signals is given as follows:

$$D = \Theta + Z \tag{4}$$

where $\Theta = [\theta_1, \ldots, \theta_M]^{\mathrm{T}}(\theta_i = [\theta_{i1}, \theta_{i2}, \ldots, \theta_{iN}])$ and Z is a column of $M \times N$ random errors with normal distribution $N(0,\sigma^2)$. The measurement error (noise) variation of the wavelet coefficients is characterized by the common process variance σ^2 for multiple signals.

3. Data classification methodology

3.1. Data pre-selection by VET

Most wavelet-related models for analyzing complicated signals have focused on feature selection and noise removal for the case of a single signal. However, many engineering applications require the simultaneous processing of multiple signals to understand the nature of a system or to extract hidden features of the defects within the system. Although a method for single-signal-based wavelet feature selection can be applied to process multiple signals, it can cause a problem in that different numbers and choices of representative wavelet features in different signals constitute no "unified wavelet-positions" in the comparison of signals, especially in the case of signals from different classes in a distinct process effect. Therefore, Jung et al. used the advantages afforded by scalograms [16] and developed the following VET procedure [13]. This wavelet feature selection procedure balances the reconstruction error against the data-reduction efficiency and proves that is powerful at capturing key patterns in multiple signals while removing the embedded noise. The selected wavelet coefficients are treated as the "reduced-size" data (reduced number of features) in subsequent analysis for decision making such as clustering and classification. The study introduced the overall relative reconstruction error (*ORRE*) function for processing multiple signals as follows:

$$ORRE(\lambda) = \Lambda(\lambda) + \xi \cdot \Upsilon(\lambda) \tag{5}$$

where

$$\Lambda(\lambda) = \frac{\sum_{m=1}^{N} E[\|d_{vm}(1 - I(\|d_{vm}\|^2 > \lambda))\|^2]}{\sum_{m=1}^{N} E[\|d_{vm}\|^2]} \tag{6}$$

$$\Upsilon(\lambda) = \frac{\sum_{m=1}^{N} E[I(\|d_{vm}\|^2 > \lambda)]}{N} \tag{7}$$

Here, E represents the expectation of random variables. Note that Eqs. (6) and (7) include the indicator function, $I(\|d_{vm}\|^2 > \lambda)$, which constitutes the threshold parameter. The indicator function is based on the "vertical energy" metric,

$$\|d_{vm}\|^2 = d_{1m}^2 + d_{2m}^2 + \cdots + d_{Mm}^2, \quad m = 1, 2, \ldots, N \tag{8}$$

which is the sum of all wavelet coefficients at the mth wavelet position. This is why it is called VET.

Further, the *ORRE* criterion was originally developed due to the requirement for balancing the reconstruction error and the data-reduction ratio. Eq. (6) represents a "normalized" reconstruction error from the wavelet approximation model $Y = W^{\mathrm{T}}D$. On the other hand, Eq. (7) indicates the number of normalized wavelet coefficients. This term is used as a penalty for including an excessive number of wavelet coefficients so that the data model can be approximated and represented in the simplest manner possible. Similar penalty ideas have been applied in ridge regression and neural networks [17]. Normally, the weighting parameter of the penalty, ξ, in Eq. (5) should be defined by the user. Alternatively, it can be provided by the generalized cross-validation (GCV) method [18]. For simplicity, this study assumes $\xi = 1$, which places equal weights on both components, Λ and Υ.

Given VET in Eq. (8), *ORRE* is minimized to determine λ. Therefore, a simple formula for estimating the optimal λ with $\lambda_{N,M}$ is developed as

$$\lambda_{NM} = \sum_{m=1}^{N} E(\|d_{vm}\|^2)/N \tag{9}$$

Since every wavelet coefficient is independent and has a normal distribution, the vertical energy of each wavelet position follows a noncentral chi-square distribution [19]. Based on this result and some other calculus derivations, Jung et al. [13] proved the optimality of λ_{NM}. Therefore, the ith position of the wavelet coefficients (across signals) can be selected if its vertical energy is larger than $\lambda_{N,M}$.

Fig. 1 shows a visual comparison of the representative original signal (Fig. 1(a)) and the reconstructed one (Fig. 1(c)) from VET wavelet positions. The original signal used in this example is simulation data from the time–history response of a cantilevered eight-bay truss model. The truss model was randomly excited and the displacement of the free end was recorded for wavelet transforms. The detailed description of the simulation process will be discussed in the following section. Here, the original signal was decomposed into eight levels using Daubechies 4 (db4), a type of wavelet. While the size of the original signal is $N = 8192$, the reconstructed one comprises only 685 VET wavelet positions, which implies a large data reduction. Figs. 1(b) and (d) are magnified versions of Figs. 1(a) and (c), respectively, that show the effectiveness of noise removal via the VET procedure. In the following analyses, the selected wavelet coefficients at the VET wavelet positions are treated as the ''size reduced'' data for developing a damage detection algorithm while achieving noise removal. Therefore, by exploiting wavelet transforms, only a small portion of the original data actually needs to be used

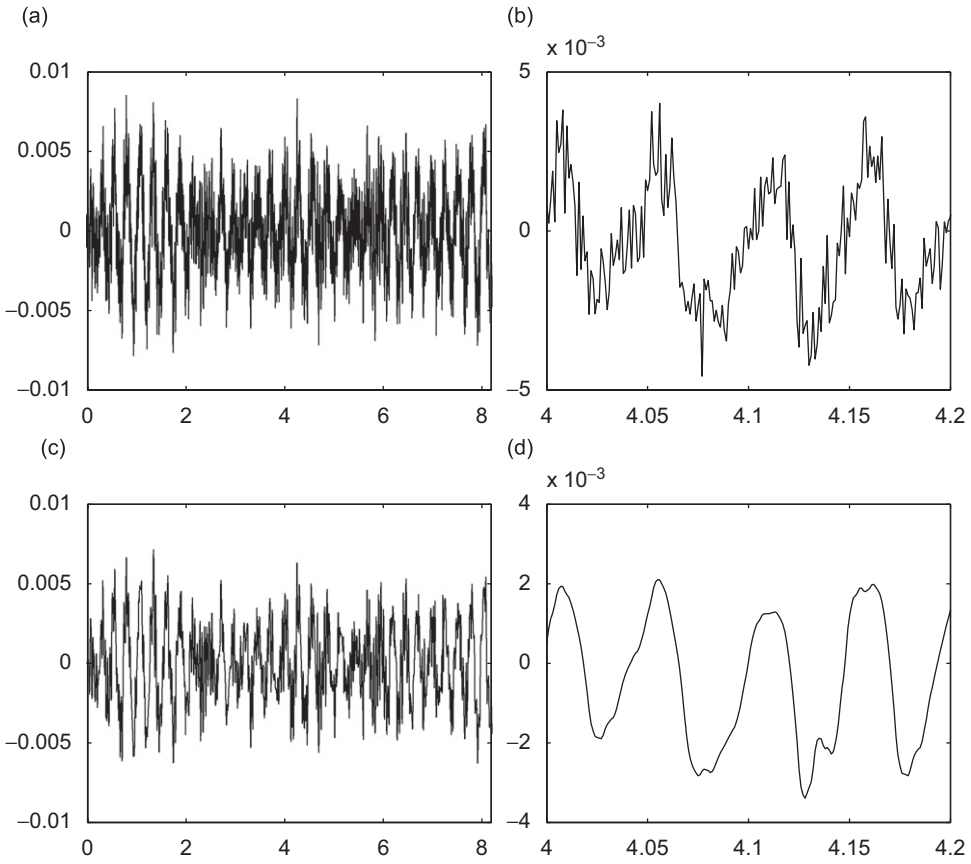

Fig. 1. Signal reconstruction using 685 VET wavelet positions and the effect of noise removal: (a) original signal, (b) magnified original signal, (c) reconstructed signal, and (d) magnified reconstructed signal.

for feature extraction, which eventually saves computing time and storage space. For more details about the VET procedure, see Jung et al. [13].

3.2. Discriminant analysis through silhouette statistics

This section describes a new approach for feature selection that generates VETS. Silhouette statistics have been widely used to assess the quality of clustering by measuring how well an object is assigned to its corresponding cluster. See Ref. [20] for more details on silhouette statistics. Here, this concept is expanded to the discriminant power function, as shown in Eq. (10) below. For signal pattern classification, it is assumed that a data set is given in which $H = (\overrightarrow{d_j}, G(j))$ for $j = 1,\ldots,M$. The data set has M data points with well-defined class labels. Note that $\overrightarrow{d_j} = (d_{1j}, d_{2j}, \ldots, d_{pj})$ is the signal vector for the jth sample described by p predictor variables that are pre-selected by VET (where $p = \sum_{m=1}^{N} I(\|d_{vm}\|^2 > \lambda_{NM})$, i.e. the number of wavelet positions selected by the VET procedure) and $G(j) \in G = \{G_1, G_2, \ldots, G_k\}$ is the class label associated with $\overrightarrow{d_j}$. Note also that k is the number of classes and n_k is the number of $\overrightarrow{d_j}$ in G_k. The proposed discriminant power function based on silhouette statistics at the ith VET feature is then defined as

$$S_i = \frac{1}{M}\sum_{j=1}^{M} \frac{b_i(\overrightarrow{d_j}) - a_i(\overrightarrow{d_j})}{\max\{a_i(\overrightarrow{d_j}), b_i(\overrightarrow{d_j})\}}, \quad i = 1, 2, \ldots, p \tag{10}$$

where, for $\overrightarrow{d_j} \in G_k$,

$$a_i(\overrightarrow{d_j}) = \frac{1}{n_k - 1} \sum_{\overrightarrow{d_{j'}} \in G_k} d_i(\overrightarrow{d_j}, \overrightarrow{d_{j'}}) \tag{11}$$

$$b_i(\overrightarrow{d_j}) = \min_{s \neq k} \frac{1}{n_k} \sum_{\overrightarrow{d_{j'}} \in G_s} d_i(\overrightarrow{d_j}, \overrightarrow{d_j}) \tag{12}$$

and

$$d_i(\overrightarrow{d_j}, \overrightarrow{d_{j'}}) = (d_{ij} - d_{ij'})^2 \tag{13}$$

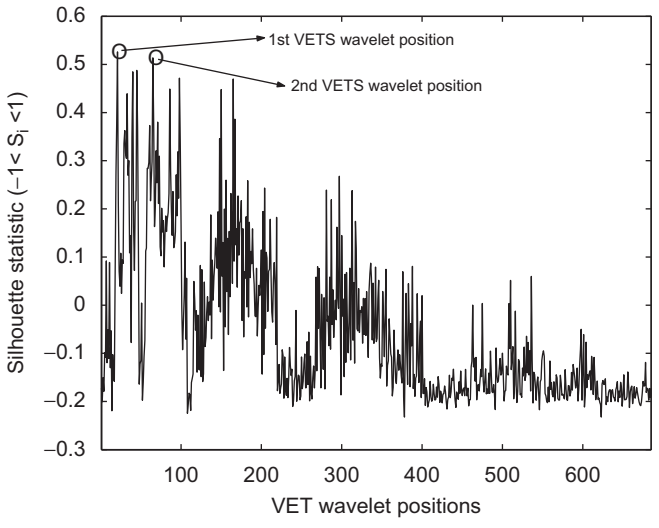

Fig. 2. Silhouette statistics of 685 VET wavelet positions.

In other words, $a_i(\overrightarrow{d_j})$ is the average distance between $\overrightarrow{d_j}$ and all other samples in the same class with respect to the ith wavelet position and $b_i(\overrightarrow{d_j})$ is the minimum average distance of $\overrightarrow{d_j}$ to all samples in other classes with respect to the ith wavelet position. The discriminant power function with respect to the ith wavelet position, S_i, returns a discriminant power score in the range of -1 to $+1$, and indicates how well all data points can be assigned to their own class in terms of the ith wavelet position. Intuitively, data points are well-classified by wavelet positions with a large silhouette statistic value, data tend to lie between classes with small silhouette values, and data points are poorly classified by those with negative values.

While relevant features can be helpful in classification, irrelevant ones may undermine the classification performance. Also, as the more features are considered, the more training examples are needed to achieve better results. Thus, choosing relevant features are critical in the process of classification. This study exploits the concept of silhouette statistic for the measure of how well the classes are discriminated with respect to a given VET feature. According to the perspective of the silhouette statistics, this study will employ S_i to select a few important wavelet positions for further cluster visualization and classification analysis. (The larger S_i of a VET feature is, the higher class-discrimination ability the VET feature has), i.e. sorting the mean silhouette statistics (discriminating power function) in ascending order:

$$S_{(1)} < S_{(2)} < \cdots < S_{(p)}, \quad p = \sum_{m=1}^{N} I(\|d_{vm}\|^2 > \lambda_{NM}) \tag{14}$$

Fig. 2 shows the silhouette statistics, S_i, $i = 1,\ldots,p$ ($p = 685$ in this simulation) for each VET wavelet position. The largest silhouette statistics S_i ($= 0.5260$) leads us to select the first VETS wavelet position (first VETS wavelet position = 38th VET wavelet position = $d_{8,6}$), the second largest S_i ($= 0.5136$) to the second position (second VETS wavelet position = 85th VET wavelet position = $d_{7,21}$), and so on. Later, when implementing CART, only five VETS containing large S_i are applied to assign the test data to their own class for evaluating the misclassification error. See Ref. [14] for more details on signal classification using CART.

4. Description of simulation

4.1. Damage in truss structure

This section describes a numerical model of a truss structure and the characteristics of simulation data. Here, the aforementioned wavelet-based classification method will be demonstrated for localizing the stiffness damage. The physical system under consideration is an eight-bay planar truss structure, as shown in Fig. 3. The truss structure is 4 m-long and has two cross-braces in each bay. All truss members comprise an

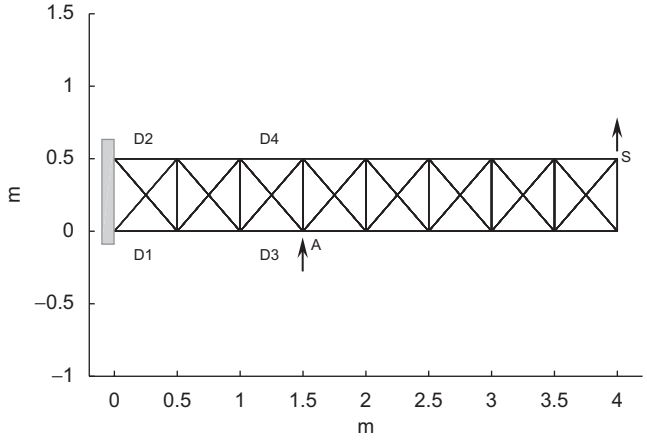

Fig. 3. Schematics of an eight-bay planar truss structure with possible damage locations ($D1$–$D4$), displacement sensor location (S), and actuator location (A).

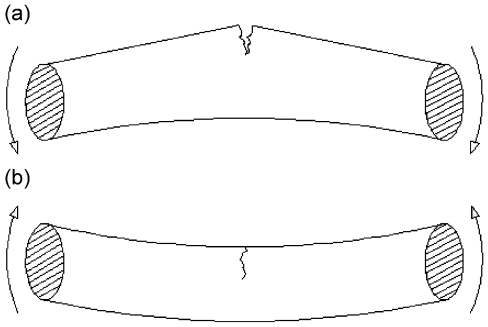

Fig. 4. Crack opening (a) and closing (b): breathing crack.

aluminium solid bar whose Young's modulus (**E**) is 70×10^9 N/m. Each strut is 2 cm in diameter and the length of each bay is 0.5 m. Its boundary condition is shown as a cantilevered truss that is fixed on a solid wall to the left. Note that a Bernoulli-beam element is used in formulating the FE model of the structure to accommodate the bending moment at each degree-of-freedom. In order to extract the dynamics of the system, Gaussian random noise input is applied as an excitation force at point A, as shown in the figure. The response of the system, i.e. the end displacement of point S along the direction of the arrow is collected as time–history data.

Two different types of damage are investigated. The first one is an open-type (slot) crack, implying that no stiffness variation occurs as the member experiences compression and expansion. In other words, the strut will have the same bending stiffness although the member undergoes compression (crack closure) and expansion (crack opening). On the other hand, a structure having a breathing-type (fatigue) crack typically behaves as a bilinear system, where the bending stiffness instantaneously changes between two states (undamaged and damaged), as shown in Fig. 4. Most damage detection researches have simply assumed their damage model to be an open-crack one in order to avoid modelling and simulation difficulties. However, a breathing-type crack is closer to a practical real-world damage mechanism. Obviously, the bilinear nature of a breathing crack largely necessitates a time–frequency data analysis technique such as wavelet transform. Therefore, this study also addresses the issue of nonlinear behaviour in breathing cracks, i.e. how they affect the overall performance of wavelet-based damage localization.

Fig. 3 also indicates four possible damage locations ($D1$–$D4$), which will be eventually localized through the proposed wavelet-based algorithm. In order to simulate structural defects or damage, the bending stiffness (**EI**) of the beam element is reduced by 50% in elements $D1$ through $D4$ (i.e. 1.0 for healthy and 0.5 for damaged state). For comparative study, the dynamic behaviour of two damage types, i.e. open- and breathing-type crack is also realized in this simulation. In order to accommodate breathing-type damage, a pair-comparison between two nodes of damaged beam elements is performed at each time step. According to the result of the current time step, i.e. the sign of the relative displacement between adjacent nodes, the elemental stiffness matrix is switched to a stiffness-reduced one for the analysis of the next time step. In other words, the compression and expansion modes of the damaged strut member alternate between different stiffness matrices in each time step. Consequently, nine scenarios are investigated, i.e. four damage location cases for two different damage types and one healthy case.

4.2. Characteristics of simulation data

The simulation generates sampled data at the rate of 1000 Hz for 20 s resulting in $N = 20{,}000$ data points. In order to facilitate uncertainties in the severity of damage, the level of reduced bending stiffness on the damaged truss member is randomly perturbed from its mean value. Therefore, the simulation is repeated 10 times to create a group of randomly populated data sets for all healthy and damaged cases, i.e. the data model of this simulation can be written as

$$y_i = \mathbf{f}(t;\beta_i) + \mathbf{e}^{(\mathbf{n})}, \quad \beta_i = g(a_i, b_i, \gamma_i, \rho_i) \tag{15}$$

where $y_i = [y_{i1}, y_{i2}, \ldots, y_{iN}]$ is a vector of N equally-spaced data points from the ith signal ($i = 1,2,\ldots,10$); $\mathbf{e^{(n)}}$, a vector of random noise; and β_i, a signal-specific parameter for the ith signal. While this variation in damage severity attempts to mimic unforeseen influences in the process error, a white noise ($\mathbf{e^{(n)}}$) is additionally imposed on each time-history data set to create measurement noise. $y_{ij} = f(t_j; \beta_i) + \varepsilon_j^{(n)}$ can be used in other expressions. In the above model, $g(\cdot)$ is an unknown function of parameters such as damage location a_i, damage severity b_i, perturbation level of damage severity γ_i, and type of damage ρ_i, i.e. open or breathing crack. The damage location parameter a_i is defined as

$$a_i \in A = \{0, 1, 2, 3, 4\} \tag{16}$$

where $a_i = 0$ indicates that the ith signal is from an undamaged (healthy) case. Similarly, if $a_i = 1,2,3$, and 4, the signals are those arising from the damaged cases at locations $D1$, $D2$, $D3$, and $D4$, respectively. b_i and γ_i are defined as

$$b_i = b + \varepsilon_i^{(b)}, \quad \varepsilon_i^{(b)} \sim N(0, (k \cdot b/3)^2) \tag{17}$$

where

$$b = \begin{cases} 1 & \text{if } a_i = 0 \\ 0.5 & \text{otherwise} \end{cases}$$

$$k = \begin{cases} 0 & \text{if } a_i = 0 \\ \gamma_i & \text{otherwise} \end{cases} \tag{18}$$

$$\gamma_i \in \Gamma = \{0.1, 0.05, 0.01\}$$

In the definition above, $\varepsilon_i^{(b)}$ is a realization of damage severity perturbation, i.e. if the signal is from a healthy case ($a_i = 0$), there is neither damage severity nor perturbation of damage severity. Unless the signal is from a healthy case, $\varepsilon_i^{(b)}$ has some value of damage severity perturbation with a parameter γ_i. Here, the value of γ_i is assumed as $Pr(-\gamma_i \cdot b < \varepsilon_i^{(b)} < \gamma_i \cdot b) = 2(\Phi(3) - 0.5) = 0.997$, where $\Phi(z) = \int_{-\infty}^{z} (1/\sqrt{2\pi}) e^{-x^2/2}\, dx$. With regard to the damage type, $\rho_i = 0$ for an open crack case while $\rho_i = f(t_j; y_{ij}(t))$ for a breathing crack, which results in the bilinear dynamic behaviour described in the previous section.

Fig. 5 shows the time–history data generated from the truss structure for the healthy case shown in Fig. 5(a) and four different damage locations, i.e. $D1$ through $D4$ in Figs. 5(b)–(e), respectively. The figures conveniently show the simulation data for the first 0.8 s. As mentioned above, all signals of each damage case include randomly perturbed damage severity (in this figure, $\gamma_i = 0.1$). The signal-to-noise ratio (SNR) is defined as $std(f)/\sigma$, where $std(f)$ is the standard deviation of the discretized signal points and σ is the standard deviation of noise. In the data model, the realization of measurement errors $\varepsilon_j^{(n)}$ is defined as

$$\varepsilon_j^{(n)} \sim N\left(0, \left(\frac{std(\mathbf{f}(\mathbf{t}; a_i = 0))}{SNR}\right)^2\right) \tag{19}$$

where $std(\mathbf{f}(\mathbf{t}; a_i = 0))$ is the standard deviation of the signal from the healthy case. Fig. 6 illustrates noise-free data (Fig. 6(a)) along with noise-corrupted ones having three levels, i.e. $SNR = 7$, 5, and 3 (Figs. 6(b)–(d)). For signals with large SNR (less noisy), the noise level (σ) is correspondingly low.

It is obvious that the proposed damage classification method should extract as many common features as possible among 10 data sets that are generated from the response of the same damage location case. This will prove the robustness of the method toward uncertainties in the perspective of damage severity. On the other hand, various damage states developed from different damage locations in the truss should be clearly distinguished to guarantee a reliable signal classification, which also tests the sensitivity of damage localization. Given the simulation data created in this study, clustering analysis will be performed to investigate the effectiveness of the VETS procedure by employing several data domains including those based on VETS wavelet positions and others. Further, this study applies CART to several domain data for classifying the damage locations and assessing its accuracy estimation. The results and discussions are presented in the next section.

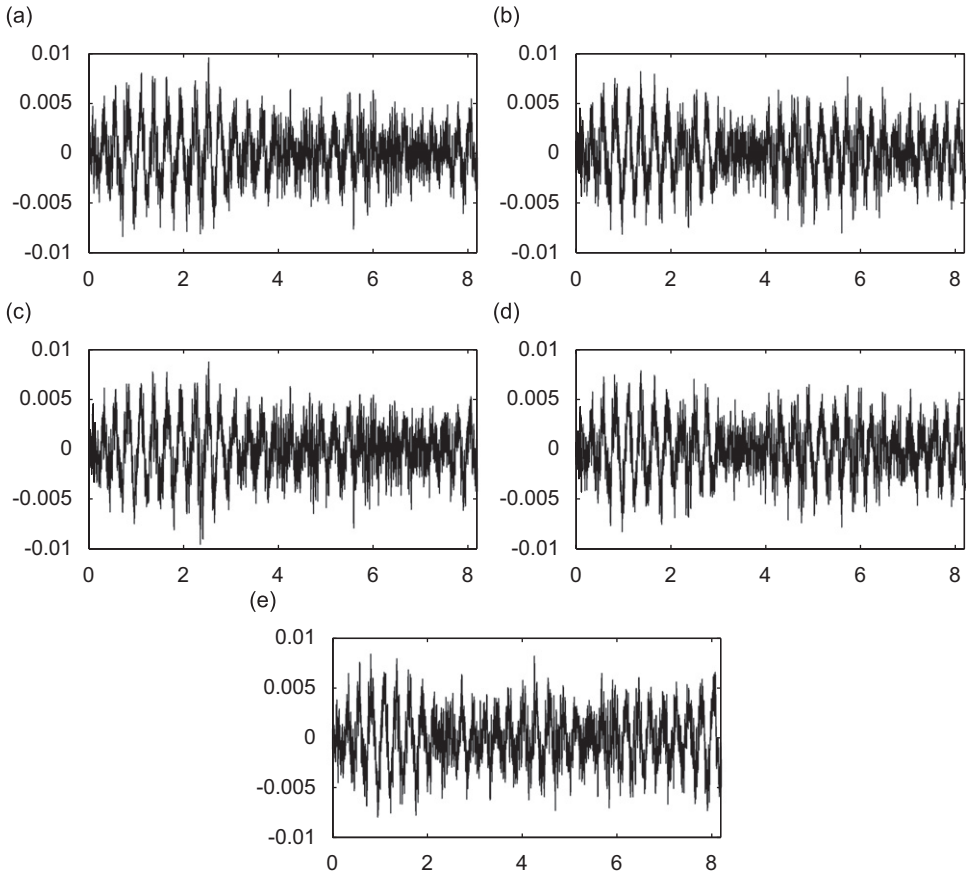

Fig. 5. Time–history response from sensor (S) in the truss structure: (a) healthy condition, (b) damage location $D1$, (c) damage location $D2$, (d) damage location $D3$, and (e) damage location $D4$.

5. Damage localization results and discussions

5.1. Clustering analysis using VETS

Simulation studies are conducted for two different damage types (open- and breathing-type cracks), three values of damage severity variation parameters ($\gamma_i \in \Gamma = \{0.1, 0.05, 0.01\}$), and three SNR levels (7, 5, and 3) on $\mathbf{e}^{(\mathbf{n})}$. The signal from a damaged case with large γ_i and small SNR inherently exhibits more variable and noisier signals in a class. Therefore, 18 different simulation cases are generated. Furthermore, these cases are created for each different damage condition, i.e. four different damage locations, respectively. Note that only three SNR cases are considered for the healthy condition.

In order to avoid many replicated 3D plots, this study shows a few representative cases with the most variable simulation setting ($\gamma = 0.1$ and $SNR = 3$) and the least variable one ($\gamma = 0.01$ and $SNR = 7$) for both open and breathing damage cases. Figs. 7–10 show some of the results from the clustering analysis for damage classification using the VETS wavelet positions. The goal of VETS-based clustering is to discriminate the existence, type, and location of damage in a truss structure. Here, the first three VETS wavelet positions are used for extracting damage-sensitive features for clustering. As shown in Fig. 7, four different damage locations, i.e. $D1$–$D4$ along with the healthy case are clearly localized in a group when they are projected to the first three VETS wavelet positions. Obviously, the damage locations in the breathing-type damage cases (Fig. 8) are less distinct than those in the open-type ones (Fig. 7) under the same perturbation parameter and SNR. This trend becomes more significant as γ increases and SNR decreases, as shown in Figs. 9 and 10. With regard to an extreme case, the largest process perturbation $\gamma = 0.1$ and the lowest $SNR = 3$ in conjunction

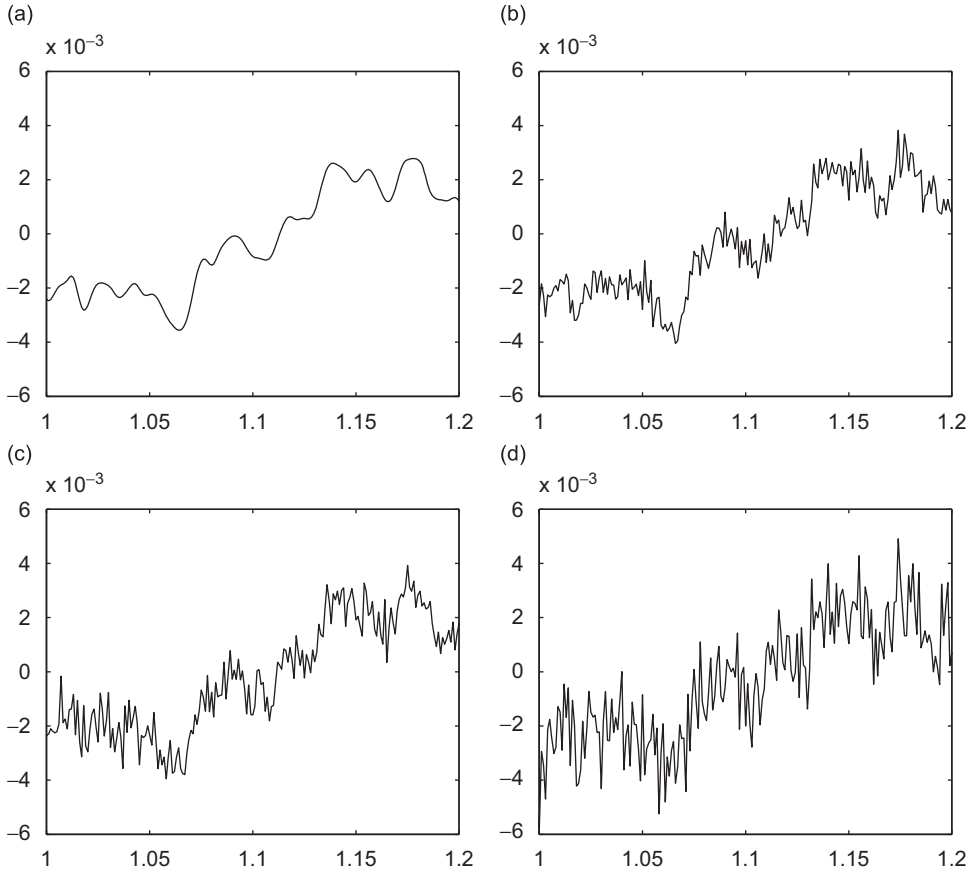

Fig. 6. Time–history responses for different *SNR*: (a) noise-free, (b) $SNR = 7$, (c) $SNR = 5$, and (d) $SNR = 3$.

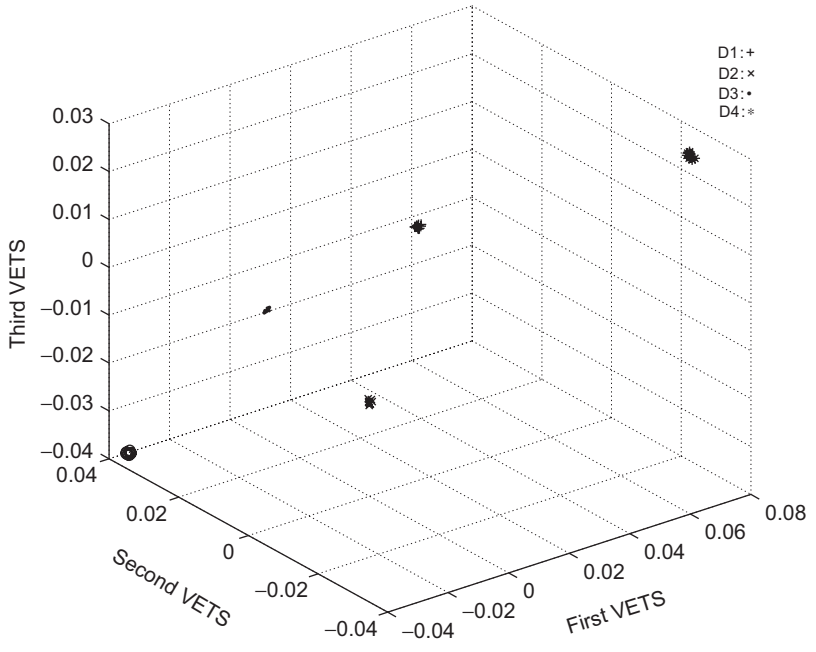

Fig. 7. VETS clustering for the open crack with $\gamma_i = 0.01$ and $SNR = 7$.

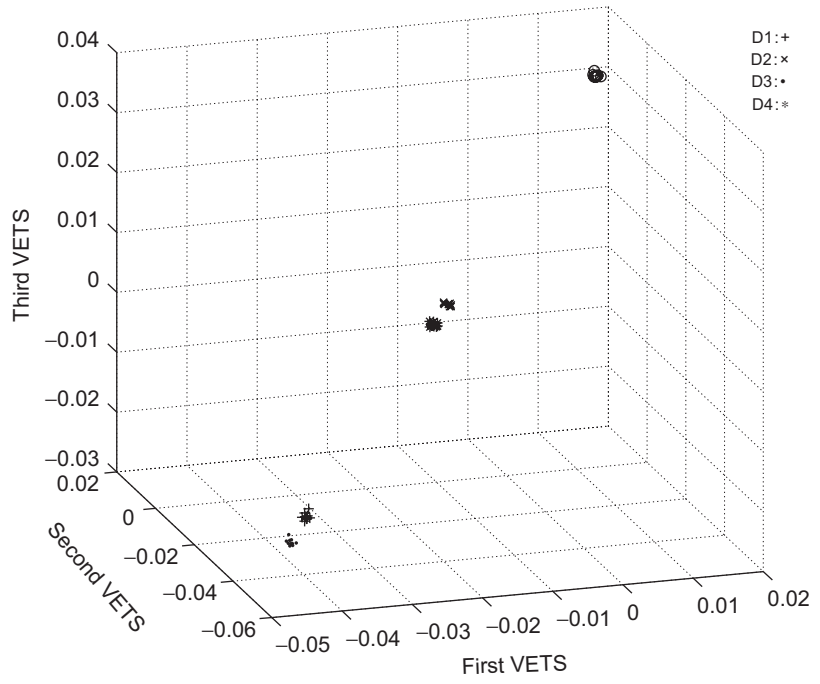

Fig. 8. VETS clustering for the breathing crack with $\gamma_i = 0.01$ and $SNR = 7$.

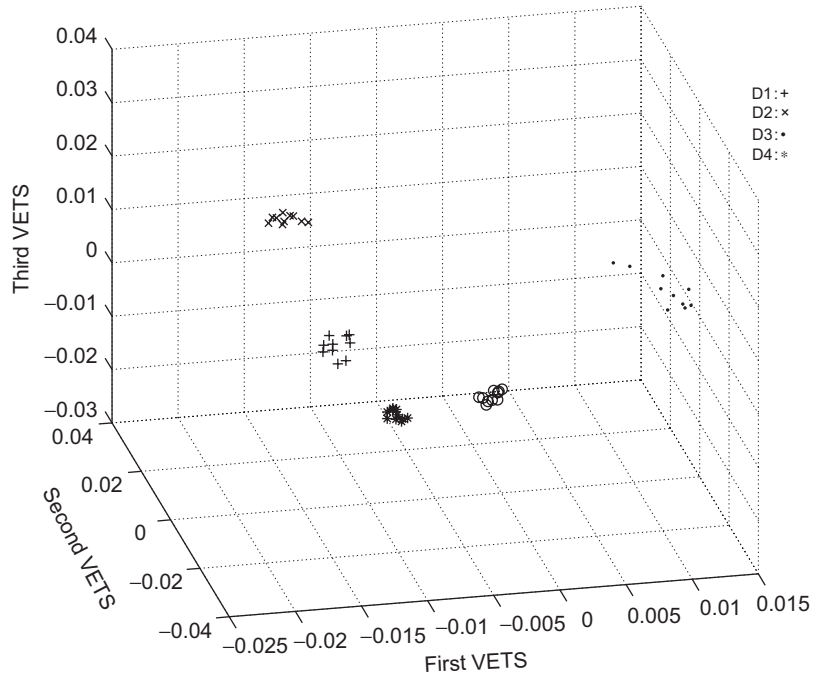

Fig. 9. VETS clustering for the open crack with $\gamma_i = 0.1$ and $SNR = 3$.

with breathing-type damage produced the worst discrimination result, as shown in Fig. 10. However, all 50 signals generated from the four different damage locations and one healthy condition (ten for each condition) have been sufficiently separated and well clustered overall for both open- and breathing-type damage. One can predict that as γ increases above 0.1, i.e. the uncertainty of damage severity increases, the size of the VETS

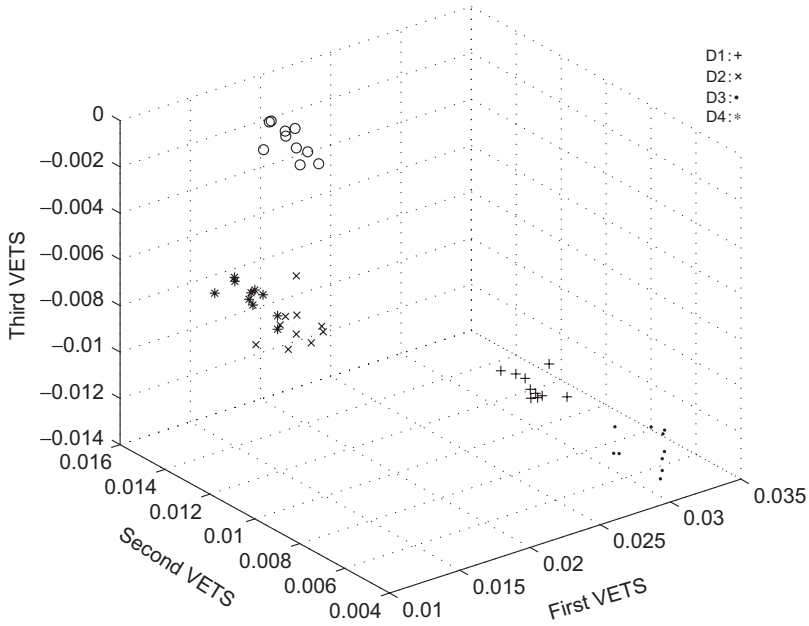

Fig. 10. VETS clustering for the breathing crack with $\gamma_i = 0.1$ and $SNR = 3$.

cluster will also increase. This trend can be used to statistically determine the confidence level of a damage location. For example, Figs. 7–10 provide a trained baseline map for confirming the presence of actual damage at unknown locations, which also exhibits the statistical level of damage severity deviation from the predefined nominal value.

5.2. Classification using CART and accuracy estimation

From the previous section, it is reasonable to assume that the central position of the VETS cluster is strongly correlated to the damage location in a given structure. Next, the CART method is applied to the data model of the VETS wavelet positions to obtain the membership index of classification for each signal. CART is a commonly used data mining tool for assessing the accuracy of classification results using the misclassification error rate as an evaluation metric. The misclassification error rate is defined as the observed proportion of misclassified observations out of the original data. Here, a ten-fold cross-validation method is adopted to obtain the misclassification error rate of CART in several data domains. In ten-fold cross-validation, a set of original data samples is partitioned into ten sub-samples. Among these sub-samples, nine are actually used for training, while the remaining sub-sample is retained as a validation data for testing the classification model, which yields the misclassification error rate. The cross-validation process is then repeated 10 times (number of folds), with each of the 10 sub-samples used exactly once as validation data. The 10 results from the folds are then averaged to yield the misclassification error rate estimation. The misclassification error rate estimations for open- and breathing-type cracks are shown in Tables 1 and 2, respectively.

By comparing these two tables, it is observed that the breathing crack has higher misclassification error rates than the open crack type, which is consistent with previous classification results. Further, CART from the VETS domain produced a more accurate classification performance as compared to other data domains. Therefore, the proposed VETS-based CART approach outperformed other methods despite the perturbed crack severity and noise-corrupted measurement error that could complicate the classification of the damage location. Note that tree-based methods such as CART partition the feature space into a set of rectangles, and then fit a simple model in each one. They are conceptually simple yet powerful. However, one major problem with tree-based method is its high variance. Often a small change in data points may result in very different

Table 1
Misclassification error rate ($\times 100 = \%$): open crack case.

γ_i	*SNR*	Time	Wavelet	VET	VETS
0.01	7	0.18	0.02	0.02	0.00
0.01	5	0.22	0.10	0.06	0.00
0.01	3	0.22	0.14	0.18	0.00
0.05	7	0.26	0.08	0.04	0.00
0.05	5	0.42	0.04	0.02	0.00
0.05	3	0.28	0.08	0.08	0.00
0.1	7	0.22	0.02	0.04	0.00
0.1	5	0.20	0.08	0.02	0.00
0.1	3	0.38	0.02	0.06	0.00

Table 2
Misclassification error rate ($\times 100 = \%$): breathing crack case.

γ_i	*SNR*	Time	Wavelet	VET	VETS
0.01	7	0.46	0.14	0.14	0.00
0.01	5	0.60	0.16	0.14	0.06
0.01	3	0.50	0.18	0.12	0.02
0.05	7	0.36	0.08	0.06	0.00
0.05	5	0.32	0.16	0.14	0.02
0.05	3	0.62	0.16	0.12	0.00
0.1	7	0.22	0.06	0.04	0.06
0.1	5	0.32	0.22	0.18	0.02
0.1	3	0.42	0.22	0.14	0.18

Table 3
Average elapsed computation time for CART.

Domain	Time	Wavelet	VET	VETS
Elapsed time (s)	198.69	196.61	16.49	0.78

series of splits, making interpretation somewhat precarious. The main reason for this instability is the hierarchical nature of the process: the effect of an error in the top split propagates down to all other the splits [17]. The worse performance of VETS than VET for the case of severity perturbation parameter = 0.1 and $SNR = 3$ in Table 1 seems due to this drawback.

The computational time of MATLAB code for CART is benchmarked using a PC with an Intel Pentium 3.2-GHz processor. Table 3 provides the comparison result of the average elapsed computation time in performing CART using data in the original time, wavelet, VET, and VETS domains. CART in the VETS domain executes faster than the others. In conclusion, the efficiency of the proposed feature extraction and selection method is validated by comparing the computation time and misclassification error.

6. Conclusions

In this paper, a novel procedure is proposed to localize structural damage by classifying the response signal in terms of VETS wavelet positions. This study uses an eight-bay planar truss structure to validate the proposed damage localization method. Two different types of damage, i.e. open and breathing conditions are investigated to test the benefits of wavelet-based signal processing. The simulated time–history responses are processed to extract damage-sensitive wavelet positions, which are further developed to classify damage

locations by exploiting VETS-based clustering analysis. The wavelet-based VET technique allows us to first obtain a suitable subset of extracted or modified features of our data, i.e. good predictor sets should contain features that are strongly correlated to the characteristics of the data without considering the classification method used, although each of these features should be as uncorrelated with each other as possible. Silhouette statistics have been used to assess the quality of clustering by measuring how well an object is assigned to its corresponding cluster. We expend this concept to the discriminant power function used in this paper. The simulation results showed that the locations of stiffness-reduced damage in the truss are successfully classified and localized even with a significant amount of noise and damage severity uncertainties. This study illustrated the potential application of wavelet-based time–frequency signal classification such as VETS in solving the damage localization problem.

References

[1] H. Sohn, C.R. Farrar, F.M. Hemez, D.D. Shunk, D.W. Stinemates, B.R. Nadler, A review of structural health monitoring literature: 1996–2001, Technical Report, LA-13976-MS, Los Alamos National Laboratory, 2003.
[2] M. Palacz, M. Krawczuk, Vibration parameters for damage detection in structures, *Journal of Sound and Vibration* 249 (5) (2002) 999–1010.
[3] A.S.J. Swamidas, Y. Chen, Monitoring crack growth through change of modal parameters, *Journal of Sound and Vibration* 186 (2) (1995) 325–343.
[4] B.H. Koh, L.R. Ray, Feedback controller design for sensitivity-based damage localization, *Journal of Sound and Vibration* 273 (1–2) (2004) 317–335.
[5] Z. Hou, M. Noori, R. St. Amand, Wavelet-based approach for structural damage detection, *Journal of Engineering Mechanics—ASCE* 12 (7) (2000) 677–683.
[6] H. Kim, H. Melhem, Damage detection of structures by wavelet analysis, *Engineering Structures* 26 (3) (2004) 347–362.
[7] W.J. Staszewski, Structural and mechanical damage detection using wavelets, *The Shock and Vibration Digest* 30 (6) (1998) 457–472.
[8] P. Moyo, J.M.W. Brownjohn, Detection of anomalous structural behaviour using wavelet analysis, *Mechanical Systems and Signal Processing* 16 (2–3) (2002) 429–445.
[9] Z. Sun, C.C. Chang, Structural damage assessment based on wavelet packet transform, *Journal of Structural Engineering—ASCE* 128 (10) (2002) 1354–1361.
[10] H. Li, X. Deng, H. Dai, Structural damage detection using the combination method of EMD and wavelet analysis, *Mechanical Systems and Signal Processing* 21 (2007) 298–306.
[11] C.-J. Lu, Y.-T. Hsu, Vibration analysis of an inhomogeneous string for damage detection by wavelet transform, *International Journal of Mechanical Sciences* 44 (2002) 745–754.
[12] B. Basu, Identification of stiffness degradation in structures using wavelet analysis, *Construction and Building Materials* 19 (2005) 713–721.
[13] U. Jung, M.K. Jeong, J.C. Lu, A vertical-energy-thresholding procedure for data reduction with multiple complex curves, *IEEE Transactions on Systems, Man, and Cybernetics—Part B* 36 (5) (2006) 1128–1138.
[14] L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, *Classification and Regression Trees*, Chapman & Hall, New York, 1984.
[15] S.G. Mallat, *A Wavelet Tour of Signal Processing*, second ed., Academic Press, London, 1999.
[16] B. Vidakovic, *Statistical Modeling by Wavelets*, Wiley, New York, 1999.
[17] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning*, Springer, Berlin, 2001.
[18] N. Weyrich, G.T. Warhola, Wavelet shrinkage and generalized cross validation for image denoising, *IEEE Transactions on Image Processing* 7 (1) (1998) 82–90.
[19] N.L. Johnson, S. Kotz, *Continuous Univariate Distributions*, Vol. 2, Houghton Mifflin, Boston, MA, 1970.
[20] P.J. Rousseeuw, Silhouettes: a graphical aid to the interpretations and validation of cluster analysis, *Journal of Computational and Applied Mathematics* 20 (1987) 53–65.