Applied Soft Computing 36 (2015) 368–374

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Structural damage detection based on posteriori probability support vector machine and Dempster–Shafer evidence theory

Qifeng Zhou[a,*], Hao Zhou[a], Qingqing Zhou[a], Fan Yang[a], Linkai Luo[a], Tao Li[b]

[a] *Department of Automation, Xiamen University, Xiamen, 361005 China*
[b] *School of Computer Science, Florida International University, Miami, FL, 33199 USA*

ARTICLE INFO

Article history:
Received 21 August 2013
Received in revised form 1 August 2014
Accepted 28 June 2015
Available online 29 July 2015

Keywords:
Damage detection
Information fusion
Support Vector Machine
DS evidence theory

ABSTRACT

An intelligent detection method is proposed in this paper to enrich the study of applying machine learning and data mining techniques to building structural damage identification. The proposed method integrates the multi-sensory data fusion and classifier ensemble to detect the location and extent of the damage. First, the wavelet package analysis is used to transform the original vibration acceleration signal into energy features. Then the posteriori probability support vector machines (PPSVM) and the Dempster–Shafer (DS) evidence theory are combined to identify the damage. Empirical study on a benchmark structure model shows that, compared with popular data mining approaches, the proposed method can provide more accurate and stable detection results. Furthermore, this paper compares the detection performance of the information fusion at different levels. The experimental analysis demonstrates that the proposed method with the fusion at the decision level can make good use of multi-sensory information and is more robust in practice.

1. Introduction

Buildings encounter structural damage of varying degrees inevitably throughout their lifetime by natural or artificial factors, which, if not handled appropriately in time, may cause immense physical destruction and enormous economic loss. Hence, how to detect the structural damage in advance has become an important issue in both academic studies and engineering applications. The damage is commonly defined as changes of the material or geometric properties of the structure system, including changes of the boundary conditions and system connectivity or the reduction of stiffness, etc. [1]. Identifying the damage accurately is then a critical task for the follow-up maintenance, repair, and rehabilitation.

Recently, data mining techniques have been widely used in structural health monitoring. One successful focus is detecting the structural damage through structural responses. The typical process of structural damage detection based on structural responses usually can be described as follows. First, a vibration monitoring system with sensors is installed inside the building and the features are extracted from the dynamic sensor signal. Second, a classifier is employed to train and predict the structural damage using the labeled samples (where the labels are the category of damage). In other words, the structural damage detection is formulated as a supervised learning problem.

A variety of machine learning and data mining methods, including Artificial Neural Network (ANN) [2,3], Genetic Algorithm [4,5], CART [6], and Support Vector Machines [7,8] are applied to damage detection. Among these methods, SVM-based methods show superior performance in generalization and accuracy.

The previous methods often use each sensor signal independently but recent research has found that information fusion (also named data fusion) is very useful to improve the identification accuracy [9]. For instance, Dempster–Shafer (DS) evidence theory, a common tool in data fusion, has been applied in detecting aircraft structure and traffic incident [10,11]. In principle, fusion of multi-sensory data provides significant advantage over single source data [12,13]. In detecting building structural damage, there are generally three kinds of level of information fusion: raw data level fusion, feature level fusion, and decision level fusion. Raw sensor data can be directly combined if the sensor data from different sources are commensurate. One example of raw data level fusion is fusing the structure vibration velocity/acceleration by cross-correlation function [14]. In feature level fusion, features are first extracted from multiple sensor observations, and then combined into a single concatenated feature vector which is input to damage detection approaches based on ANN, SVM or Random Forest [3,15]. Finally, decision level fusion involves semantic fusion of sensor information, after each sensor has made a preliminary determination of a structural damage location and extent.

* Corresponding author. Tel.: +86 592 2580036; fax: +86 592 2580102.
E-mail address: zhouqf@xmu.edu.cn (Q. Zhou).

http://dx.doi.org/10.1016/j.asoc.2015.06.057
1568-4946/

In practice, a fusion of sensor data may actually produce worse results than those could be obtained by tasking the most appropriate sensor in a single source. This is often due to the attempt to combine accurate (i.e., good data) with inaccurate or biased data, especially if the uncertainties or variances of the data are unknown [16]. Once the inaccurate data is fused (for example, some sensors failure or installed in an inappropriate place), there is no simple way to eliminate the bad effect in the first two fusion approaches (e.g., raw data level fusion and feature level fusion). Furthermore, how to choose the appropriate algorithms or techniques to improve the stability of identification are also the key problems in structural damage detection.

In this paper, we propose an improved SVM-based detection method to increase the applicability and precision of building structural damage detection. Moreover, the augmented SVM will be combined with the DS evidence theory to provide a decision level information fusion. Our experimental study demonstrated that the decision level fusion can provide a more accurate and stable performance than commonly used SVM-based method.

The remaining of the paper is organized as follows. Section 2 introduces the wavelet package decomposition (WPD) which is used in the feature extraction from the sensor signals. Subsequently, we introduce the basic idea of the posterior probability SVM and the DS evidence theory. The core of our proposed method with the detailed description of the application is illustrated in Section 4. The experiment is shown in Section 5 and we conclude in Section 6.

2. The wavelet package decomposition (WPD)

WPD is an extension of the discrete wavelet transform by a generalization of the link between multiresolution approximation and wavelets, where, in addition to dividing the approximations of the signal, the details are also divided [17].

For N levels of decomposition, WPD produces 2^N different band signals D_{Nj}, ($j=0, 1, 2, ..., 2^N-1$) from the low frequency to the high frequency. The energy for each frequency band of signals is calculated as Eq. (1).

$$E_{Nj}=\int |D_{Nj}(t)|^2dt=\sum_{k=1}^{n}|d_{jk}|^2 \tag{1}$$

where d_{jk} is the kth discrete point's amplitude of the reconstructed signal D_{Nj} and n is the number of discrete points.

According to Eq. (1), the energy for each frequency band is first calculated and then used as elements to construct a feature vector S_N

$$S_N=[E_{N0},E_{N1},\ldots,E_{Nj},\ldots,E_{N(2^N-1)}] \tag{2}$$

Normalizing the above feature vectors by column, the new feature vector S'_N can be obtained as follows:

$$S'_N=[E'_{N0},E'_{N1},\ldots,E'_{Nj},\ldots,E'_{N(2^N-1)}] \tag{3}$$

3. Posterior probability SVM (PPSVM) and DS evidence theory

3.1. Posterior probability SVM

The standard SVM in classification cannot give the posterior probability of the sample which may be very useful in uncertainty problems. Platt [22] proposed utilizing the Sigmoid Function to map the outcome of standard SVM to a probabilistic value, as shown in Eq. (4):

$$P(y=1|f(x))=\frac{1}{1+\exp(Af(x)+B)}, \tag{4}$$

where $f(x)$ is the output of standard SVM, A and B are two parameters that can be attained by solving the following optimization problem:

$$\min_{z=(A,B)^T}\left\{-\sum_{i=1}^{l}[t_i\log(p_i)+(1-t_i)\log(1-p_i)]\right\} \tag{5}$$

where

$$p_i=\frac{1}{1+\exp(Af(x_i)+B)}$$

$$\text{and}\quad t_i=\begin{cases}\dfrac{N_++1}{N_++2} & \text{if}\, y_i=1\\[2mm] \dfrac{1}{N_-+2} & \text{if}\, y_i=-1\end{cases},\quad i=1,2,\ldots,l,$$

in which N_+ and N_- are the number of the positive and negative samples, respectively.

With the modification of the output of standard SVM, more information about the sample could be preserved which plays an important role in the subsequent processing.

3.2. DS evidence theory

The DS evidence theory is a kind of reasoning and processing method handling uncertainty problems [18,19]. It combines evidence from different sources and arrives at a degree of belief (represented by a belief function) that takes into account all the available evidence.

3.2.1. Basic concept

Assuming the possible solutions to a judgment problem constitute a set, denoted as Θ, which is also named as the framework of identification. Given the evidence, we can get a belief function on the framework, reflecting the true value assigned by the evidence of the possible proposition on the framework. Assuming that there are multiple kinds of evidence acting on the framework, which leads to multiple belief functions, we can utilize the Dempster Fusion Rule to combine these results to generate a unique belief function. The result then can illustrate the probability of the true value in some set.

Definition 1 *(:)*. Basic Probability Assignment (BPA)

Assuming Θ denotes a set containing all possible values of X, it is the framework of identification of X, and the elements in Θ are incompatible. Define a function $m: 2^{\Theta}\to[0,1]$ meeting the following conditions:

$$\text{(i)}\, m(\Phi)=0;\quad \text{(ii)}\sum_{A\in 2^{\Theta}}m(A)=1;$$

Then m is the basic probability assignment, and $m(A)$ is the basic probability value of A. When $A\neq\Theta$, $m(A)$ represents the precise level of trust and when $A=\Theta$, $m(\Theta)$ represents the assignment to the unknown.

Definition 2 *(:)*. Belief Function

The belief function of the proposition A is defined as

$$Bel(A)=\sum_{B\subseteq A}m(B),\quad \forall A\subseteq\Theta \tag{6}$$

It is also named as the lower bound, meaning the least probability that the proposition holds. If a number of evidence supports one proposition, then it should also support the inference of the proposition similarly. Hence, the belief on one proposition equals the sum of the belief on its all premises by evidence. From the definition of BPA, we know $Bel(\Phi)=0$, $Bel(\Theta)=1$.