Mechanical Systems and Signal Processing 105 (2018) 1–15

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

A comparison of linear approaches to filter out environmental effects in structural health monitoring

A. Deraemaeker [a,*], K. Worden [b]

[a] Université Libre de Bruxelles (ULB) – BATir, 50 av Franklin Roosevelt, CP 194/02, 1050 Brussels, Belgium
[b] Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK

ARTICLE INFO

Article history:
Received 7 September 2017
Received in revised form 22 November 2017
Accepted 30 November 2017

Keywords:
Structural Health Monitoring (SHM)
Environmental effects
Mahalanobis squared-distance
Principal Component Analysis (PCA)
Factor analysis

ABSTRACT

This paper discusses the possibility of using the Mahalanobis squared-distance to perform robust novelty detection in the presence of important environmental variability in a multivariate feature vector. By performing an eigenvalue decomposition of the covariance matrix used to compute that distance, it is shown that the Mahalanobis squared-distance can be written as the sum of independent terms which result from a transformation from the feature vector space to a space of independent variables. In general, especially when the size of the features vector is large, there are dominant eigenvalues and eigenvectors associated with the covariance matrix, so that a set of principal components can be defined. Because the associated eigenvalues are high, their contribution to the Mahalanobis squared-distance is low, while the contribution of the other components is high due to the low value of the associated eigenvalues. This analysis shows that the Mahalanobis distance naturally filters out the variability in the training data. This property can be used to remove the effect of the environment in damage detection, in much the same way as two other established techniques, principal component analysis and factor analysis. The three techniques are compared here using real experimental data from a wooden bridge for which the feature vector consists in eigenfrequencies and modeshapes collected under changing environmental conditions, as well as damaged conditions simulated with an added mass. The results confirm the similarity between the three techniques and the ability to filter out environmental effects, while keeping a high sensitivity to structural changes. The results also show that even after filtering out the environmental effects, the normality assumption cannot be made for the residual feature vector. An alternative is demonstrated here based on extreme value statistics which results in a much better threshold which avoids false positives in the training data, while allowing detection of all damaged cases.

1. Introduction

Vibration-based Structural Health Monitoring (SHM) techniques have been around for many years and are still today an active topic of research. Despite this fact, very few industrial applications exist. Two major trends coexist in the field: model-based and data-based techniques. Model-based techniques are often sophisticated and require a high degree of engineering

---

* Corresponding author.
E-mail address: aderaema@ulb.ac.be (A. Deraemaeker).

https://doi.org/10.1016/j.ymssp.2017.11.045
0888-3270/

knowledge and heavy hardware and software resources; they have however more potential to cover all levels of SHM, from damage detection to damage prognosis. On the other hand, data-based techniques are appealing as they require less engineering knowledge as well as limited hardware and software resources. From that point of view, they are ideal candidates for industrial applications. These methods are however generally limited to the lowest levels of SHM: damage detection and in some cases, damage localisation or classification and severity if training data are available [1].

Data-based damage detection techniques consist in detecting a deviation from the normal condition based only on undamaged-state data measured on the structure or system to be monitored. This paper focuses on the use of vibration data collected at regular time intervals. A further step consists in feature extraction, i.e. transforming these time series data into meaningful information, called features (the most common being the mode shapes and eigenfrequencies). The stochastic nature of the excitation and the unavoidable added noise on the measured data results in features having a stochastic nature. This fact being recognised, it is natural to turn to statistical methods to monitor them and detect any significant deviation from the normal condition.

The three basic elements of data-based damage detection are therefore (i) a permanent sensor network system, (ii) an automated procedure for real-time or periodic feature extraction, and (iii) a robust novelty detector [2]. The first element has received much attention in the last decade and the enormous advances in sensors and instrumentation make it possible today to deploy very large sensor networks on structures and gather the measured data in central recording units at high sampling rates. The second element is still today a challenge, and for the most widely-used features (eigenfrequencies and mode shapes) is an active topic of research [3,4]. An alternative is to look at other features which can easily be extracted automatically from the time domain data. Several efforts have been made in that direction, such as the use of residuals based on Hankel matrices [5], or peak indicators in the frequency output of modal filters [3]. For the third element, different approaches have been borrowed from statistics and machine learning, the most commonly used being control charts [6], outlier analysis using the Mahalanobis squared-distance [7] (which is similar to Hoteling T^2 control charts for individual measurements) or hypothesis testing [5].

This paper focuses on the third element of the data-based damage detection system in the presence of confounding effects due to environmental and operational variability. Important variability of the dynamic properties of structures under ambient vibrations has been identified in a number of studies [8–11]. These studies highlight the fact that variations in the dynamic characteristics due to confounding effects (temperature, humidity, traffic, wind) can be of the same order of magnitude as, or greater than, the variations due to damage, which hinders detection of the onset of damage. In the context of data-based damage detection, which is the focus of this paper, several methods have been proposed in order to filter out these effects. The simplest methods are based on the identification of the linear subspace to which the environmental and operational conditions belong, in order to remove their effect on the monitored features. Such methods are well suited when the dimension of the feature vector is large enough to be able to find a linear subspace to which the confounding effects belong. When such is not the case, non-linear methods which consist in identifying a non-linear manifold instead of a linear subspace can be used [12–14]. These techniques are however much more complex and computationally costly. The present paper has two aims: (i) to show that the Mahalanobis squared-distance can be used to filter out confounding effects in a very similar way to the linear techniques, and (ii) to demonstrate the link between the aforementioned technique and the two most commonly used linear techniques to filter confounding effects: principal component analysis [15] and factor analysis [16].

This paper is organised as follows: the second section details the mathematical foundations of the methodology proposed to filter out confounding effects based on the Mahalanobis squared-distance. The third section presents briefly the two most commonly-used linear techniques for filtering out confounding effects: principal component analysis and factor analysis, and shows the link with the method proposed in Section 2. In the fourth section, the three techniques are applied on data acquired from a wooden bridge under changing environmental conditions. The results clearly demonstrate the strong similarities between the three techniques.

2. Multivariate novelty detection under changing environmental conditions

2.1. The Mahalanobis squared-distance

Consider a set of N feature vectors $\{y_i\}$ $(i = 1 \dots N)$ of dimension n, representing N samples of the 'healthy state' of a structure, of which the mean vector $\{\bar{y}\}$ of size $n \times 1$ and the covariance matrix $[C]$ of size $n \times n$ can be estimated as follows,

$$\{\bar{y}\} = \frac{1}{N}\sum_{i=1}^{N}\{y_i\} \tag{1}$$

$$[C] = \frac{1}{N-1}\sum_{i=1}^{N}(\{y_i\} - \{\bar{y}\})(\{y_i\} - \{\bar{y}\})^T \tag{2}$$

The multivariate feature vectors correspond to the features extracted from the vibration measurements such as a set of eigenfrequencies, modeshapes, FRF or transmissibility functions at given frequencies, etc. (Throughout this paper, curved braces denote vectors, and square braces denote matrices).