Mechanical Systems and Signal Processing 50-51 (2015) 467–479

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Applying robust variant of Principal Component Analysis as a damage detector in the presence of outliers

Fahit Gharibnezhad*, Luis E. Mujica, José Rodellar

Control, Dynamic and Applications Group (CoDAlab), Department of Applied Mathematics III, Escola Universitària d'Enginyeria Tècnica Industrial de Barcelona (EUETIB), Universitat Politècnica de Catalunya (UPC), Comte Urgell 187, 08036 Barcelona, Spain

A R T I C L E I N F O

Article history:
Received 22 October 2013
Received in revised form
22 March 2014
Accepted 24 May 2014
Available online 17 June 2014

Keywords:
Damage detection
Robust PCA
Outliers

A B S T R A C T

Using Principal Component Analysis (PCA) for Structural Health Monitoring (SHM) has received considerable attention over the past few years. PCA has been used not only as a direct method to identify, classify and localize damages but also as a significant primary step for other methods. Despite several positive specifications that PCA conveys, it is very sensitive to outliers. Outliers are anomalous observations that can affect the variance and the covariance as vital parts of PCA method. Therefore, the results based on PCA in the presence of outliers are not fully satisfactory. As a main contribution, this work suggests the use of robust variant of PCA not sensitive to outliers, as an effective way to deal with this problem in SHM field. In addition, the robust PCA is compared with the classical PCA in the sense of detecting probable damages. The comparison between the results shows that robust PCA can distinguish the damages much better than using classical one, and even in many cases allows the detection where classic PCA is not able to discern between damaged and non-damaged structures. Moreover, different types of robust PCA are compared with each other as well as with classical counterpart in the term of damage detection. All the results are obtained through experiments with an aircraft turbine blade using piezoelectric transducers as sensors and actuators and adding simulated damages.

1. Introduction

Principal Component Analysis (PCA) plays a vital role in statistical analysis. PCA, also known as Karhunen–Loève decomposition or Proper Orthogonal Decomposition (POD), is a multivariate statistical technique that was first introduced by Pearson [1], developed independently by Hotelling [2], and used for first time in the mechanics community by Lumley [3].

The central idea of principal component analysis is to reduce the dimensionality of a data set in which there are a large number of interrelated variables, while retaining as much as possible of the variation present in the data set [4].

Dealing with high dimensional data, the first step in the data analysis is a dimensionality reduction. There are different reasons for that; for instance, the multidimensional data sets are difficult to interpret, and their structure cannot be visualized directly. In addition, the redundant variables create empty space and computational problems. PCA is probably the most useful tool to solve these problems [4]. Classical PCA is a linear transformation that maps the data into a lower dimensional space by preserving as much data variance as possible. In other words, it transforms the data to a new

* Corresponding author. Tel.: +34 610192812.
E-mail addresses: fahit.gharibnezhad@upc.edu (F. Gharibnezhad), luis.eduardo.mujica@upc.edu (L.E. Mujica), jose.rodellar@upc.edu (J. Rodellar).

http://dx.doi.org/10.1016/j.ymssp.2014.05.032
0888-3270/

coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on [1]. If the data compression is sufficient, the large number of variables is substituted by a small number of uncorrelated latent factors which can explain sufficiently the data structure. In addition to its main usage as a dimension reduction tool, it can be used as pattern recognition, data interpretation, visualization, outlier detection and clustering tool in various engineering fields including Structural Health Monitoring. In next few paragraphes, we will review the latest application of PCA in SHM. After that, we will consider the drawbacks of PCA and we will mention the demand to reconsider the SHM techniques that are using PCA.

PCA has been widely used in Structural Health Monitoring (SHM) in different ways [5–9]. A brief general review of some applications is presented as follows. Johnson [10] uses PCA as a clustering tool for acoustic emission (AE) transients in a composite laminate to detect matrix cracking and delamination. Manson et al. [11] apply PCA as visualization and clustering tool on acoustic emission data from box girder of a bridge for tracking crack development. Li et al. [12] use PCA as dimension reduction, classifier and visualization method for monitoring different operating conditions of automobile gearbox. Tong et al. [13] apply PCA on power spectral density (PSD) of impact-acoustic data to inspect defects assessment of tile-wile. Also they use a first two principal components as feature vectors of an artificial neural network (ANN) for classification purposes. Mustapha et al. [14] apply PCA on transmitted ultrasonic guided wave in hollow cylinder-like structure to classify undamaged conditions from damaged one. In another work, the same approach is applied on aircraft wing holder [15]. Cammarata et al. [16] use PCA beside wavelet transform for fatigue crack detection. Golinval et al. [17] and De Boe and Golinval [18] define a damage index as a comparison between the hyperplanes (subspace spanned by PCA) associated with healthy and damaged structures. Pirra et al. [19] use PCA as a monitoring technique for the diagnosis of bearing faults. In their works, they consider the effects of environmental condition changes as well as proposing a novelty criterion to isolate the area in which the faulty bearing stands. Zimroz and Bartkowiak [20] use PCA to find out the intrinsic dimension of data in both “bad” and “good” conditions and claim that intrinsic dimension of data could be used as a damage index in machine condition monitoring. He et al. [21] use PCA as a machine conditioning tool on internal-combustion engine as well as automobile gearbox. Moreover, they propose a method to select the most appropriate representative PCs somehow that the selected low-dimensional PC representations respect the best statistical information of the machine. Mujica et al. [22] implement T^2 and Q statistic on guided waves for damage detection. Their technique is based on the creation of a statistical data-driven PCA model for normal operation condition, so that any fault condition causes deviation from this model that can be detected using statistic indices. Baydar et al. [23] apply the same approach to detect, localize and identify the severity of very early tooth defects (deformation, breakage or fracture of the teeth) on industrial helical gearbox. Lili [24] implements PCA on hardware to establish a distributed system to achieve the goal of on-line monitoring and real-time damage detection.

In a specific case of SHM, damage detection and localization, the way that PCA is used can be divided into two modes. In the first mode it plays the direct role to detect and localize the damages and in the second mode, PCA is used as a primary step for other algorithms. The works in Zimroz and Bartkowiak [20] and Pirra et al. [19] are some examples that belong to the first type. As more examples, Mujica et al. [25] first apply PCA to remove the effect of environmental changes and operational condition and then use residual errors of PCA as a damage indicator for long term structural health monitoring. Besides, Tondreau et al. [26] apply PCA on frequency response function as well as transmissibility for an experimental damage detection of aircraft wing. The work in Mujica et al. [22] is an example of the second type. As more examples, Tibaduiza et al. [27] implement two indices to detect and localize damages based on a model obtained from projecting information using PCA. Gharibnezhad et al. [28] use Andrew plots as new index based on PCA for damage detection and classification in a commercial aircraft turbine blade.

However, despite the mentioned features, PCA is known to possess some shortcomings. One of them is the effect that anomalous data have on its efficiency. As both the classical variance (which is being maximized) and the classical covariance matrix (which is being decomposed) are very sensitive to outlying observations, the first components are often attracted towards outlying points, and may not capture the variation of the regular observations. Therefore based on the importance and variety of PCA usage in SHM field such as data reduction, modeling or any other algorithms that directly or indirectly involves PCA, all of them may become unreliable if outliers are present in the data.

Outliers are generally inevitable due to different reasons such as experimental fluctuation. In the best situation, one can recognize outliers on PCA's plot and remove them and repeat the algorithm again but this is not the optimized solution as it is time consuming. In addition, in many cases, it may happen that one outlier is visible but at the same time masks all the others, or several outliers can act together in such a way to diminish or even cancel each other's influence. To deal with this case more efficiently, a better way is to apply a robust variant of PCA.

To deal with the mentioned problem, in this work damage detection and classification based on different Robust PCA methods are presented. The main novelty of this work is proposing and evaluating the strategy of using robust PCA in SHM field instead of classical version. To do that, the results of using robust PCA are compared with the classical counterpart. This work shows that robust PCA could be used instead of the classical one in each of mentioned category of PCA usage that leads to more reliable outcome. Moreover, here we compare different robust PCA with each other and evaluate the efficiency of each method in damage detection. To achieve these goals, we apply PCA directly on the original data captured from the structure for damage detection goals. According to the result, damage detection based on robust PCA is more successful and, in many cases, allows the detection where classical PCA is not able to discern between damaged and non-damaged structures.