Mechanical Systems and Signal Processing 92 (2017) 196–212

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Agglomerative concentric hypersphere clustering applied to structural damage detection

Moisés Silva [a,*], Adam Santos [a], Reginaldo Santos [a], Eloi Figueiredo [b,c], Claudomiro Sales [a], João C.W.A. Costa [a]

[a] Applied Electromagnetism Laboratory, Universidade Federal do Pará, R. Augusto Corrêa, Guamá 01, Belém, 66075-110 Pará, Brazil
[b] Faculty of Engineering, Universidade Lusófona de Humanidades e Tecnologias, Campo Grande 376, 1749-024 Lisbon, Portugal
[c] CONSTRUCT, Institute of R&D in Structures and Construction, R. Dr. Roberto Frias s/n, 4200-465 Porto, Portugal

ARTICLE INFO

Article history:
Received 31 August 2016
Received in revised form 15 January 2017
Accepted 18 January 2017

Keywords:
Agglomerative concentric hypersphere
Clustering
Damage detection
Structural health monitoring
Environmental conditions
Operational conditions

ABSTRACT

The present paper proposes a novel cluster-based method, named as agglomerative concentric hypersphere (ACH), to detect structural damage in engineering structures. Continuous structural monitoring systems often require unsupervised approaches to automatically infer the health condition of a structure. However, when a structure is under linear and nonlinear effects caused by environmental and operational variability, data normalization procedures are also required to overcome these effects. The proposed approach aims, through a straightforward clustering procedure, to discover automatically the optimal number of clusters, representing the main state conditions of a structural system. Three initialization procedures are introduced to evaluate the impact of deterministic and stochastic initializations on the performance of this approach. The ACH is compared to state-of-the-art approaches, based on Gaussian mixture models and Mahalanobis squared distance, on standard data sets from a post-tensioned bridge located in Switzerland: the Z-24 Bridge. The proposed approach demonstrates more efficiency in modeling the normal condition of the structure and its corresponding main clusters. Furthermore, it reveals a better classification performance than the alternative ones in terms of false-positive and false-negative indications of damage, demonstrating a promising applicability in real-world structural health monitoring scenarios.

1. Introduction

In the last few decades, the continuous structural condition assessment has demanded strong research efforts to support the management of structures during their lifetime. In particular, for the civil engineering infrastructure [1], the bridge management systems (BMSs) and structural health monitoring (SHM) have been used to cover the most relevant activities concerning the bridge management process. The BMS is a visual inspection-based decision-support tool developed to analyze engineering and economic factors and to assist the authorities in determining how and when to make decisions regarding maintenance, repair and rehabilitation of structures [2,3]. On the other hand, the SHM traditionally refers to the process of implementing monitoring systems to measure structural responses in real-time and to identify anomalies and/or damage at early stages [4].

* Corresponding author.
E-mail address: moises.silva@icen.ufpa.br (M. Silva).

http://dx.doi.org/10.1016/j.ymssp.2017.01.024
0888-3270/

Even with the inherent limitation imposed by the visual inspections, the BMS has already been accepted by structural managers around the world [5–7]. At the same time, SHM is becoming increasingly attractive due to its potential ability to detect damage, contributing positively to life-safety and economical issues [8]. It can be also integrated into the BMS in a systematic way [9]. Posed in the context of a statistical pattern recognition (SPR) paradigm, the SHM is described as a four-phase process [10]: (1) operational evaluation, (2) data acquisition, (3) feature extraction, and (4) statistical modeling for feature classification.

The feature extraction phase estimates damage-sensitive features, from the raw data, which are potentially correlated to the level of damage present in the monitored structure. Nevertheless, when one deals with real-world monitoring scenarios, the influence of operational and environmental effects may cause changes in the magnitude of features as well as alter their correlation with the level of damage. Generally, the more sensitive a feature is to damage, the more sensitive it is to changes in the operational and environmental conditions (e.g. temperature, wind speed and traffic loading) [11]. Therefore, robust feature extraction methods and data normalization procedures are required to overcome this problem. The data normalization is the process of separating changes in damage-sensitive features caused by damage from those caused by varying operational and environmental conditions [12,13]. These influences on the structural response have been cited as one of the major challenges for the transition of SHM technology from research to practice [14–17]. Although the data normalization occurs in almost all phases (except the first one) of the SPR paradigm, the focus of this study is on the fourth phase, which is concerned with implementation of algorithms to analyze and learn the distributions of the extracted features, aiming to distinguish between the undamaged and damaged structural state conditions [18].

Some studies in literature have established the concept of automatically discovering and characterizing the normal condition of structures as clusters [9,19]. In those studies, the damage detection strategy is carried out as an outlier detection approach based on machine learning algorithms and distance metrics, allowing one to track the outlier formation in relation to time. One important note is related to the output-only nature of those approaches, implying the accomplishment of data normalization without any information about the operational or environmental parameters.

Highlighting the fact that in most engineering infrastructure only data from undamaged (or normal) condition is available on the training phase, the unsupervised learning algorithms are often required for data normalization purposes [20], i.e. training is carried out using only data from the normal structural condition. In this context, cluster-based algorithms are attractive due to their ability to discover groups of similar observations related to the same structural state at a given period of time. This unsupervised implementation permits one to detect damage formation regarding the chosen main groups of states [21–23]. Although numerous traditional unsupervised machine learning algorithms have been reported [24–27], herein the approaches based on Mahalanobis Squared Distance (MSD) and Gaussian Mixture Model (GMM) are relevant, due to their cluster-based performance, operating in a set of stable and undamaged state conditions [9,19].

In this paper, a straightforward and nonparametric method based on agglomerative clustering and inflated hyperspheres is proposed to learn the normal condition of structures. The proposed method does not require any input parameter, except the training data matrix. Two deterministic initialization procedures rooted on eigenvectors/eigenvalues decomposition and uniform data sampling are presented. Furthermore, a random initialization is also introduced. These mechanisms pave the way for a novel Agglomerative Concentric Hypersphere (ACH) algorithm that discovers an appropriate number of clusters using a density-based approach.

The classification performance is investigated on the basis of Type I/Type II errors (false-positive and false-negative indications of damage, respectively) trade-off through application on two data sets from the Z-24 Bridge, located in Switzerland. The remainder of this paper is organized as follows. In Section 2, a review of the most traditional machine learning and cluster-based approaches for structural damage detection is introduced. The clustering constraints and initialization procedures related to the ACH algorithm are presented in Section 3. Section 4 is devoted to describe the data sets used as damage-sensitive features from the Z-24 Bridge. Section 5 describes the experimental results and carries out comparisons and discussions. This study concludes in Section 6 with a summary and the main conclusions.

2. Related work

Traditionally, in most civil applications, the damage detection process is carried out using physics-based methods and parametric approaches. However, in complex engineering structures, those methods may not be practical due to the level of expertise and time required to their development [28,29]. On the other hand, nonparametric approaches rooted in the machine learning field, especially cluster-based ones, have become an alternative, as they are very useful to find hidden patterns from monitoring data and are computationally efficient [30,31]. Herein, machine learning-based approaches addressing damage assessment are discussed; moreover, the most relevant cluster-based methods and their adaptation to damage detection in SHM are also considered.

2.1. Machine learning approaches for damage detection

Principal component analysis (PCA) is a common method to perform data normalization and feature classification without measurements of the sources of variability. Yan et al. [32] presented a PCA-based approach to model linear environmental and operational influences using only undamaged features. The number of principal components from extracted features