Steel and Composite Structures, *Vol. 29, No. 6 (2018) 703-716*
DOI: https://doi.org/10.12989/scs.2018.29.6.703

An intelligent health monitoring method for processing data collected from the sensor network of structure

Ramin Ghiasi [a] and Mohammad Reza Ghasemi *

Department of Civil Engineering, Faculty of Engineering, University of Sistan and Baluchestan, Zahedan, Iran

(*Received February 11, 2018, Revised August 20, 2018, Accepted December 21, 2018*)

Abstract. Rapid detection of damages in civil engineering structures, in order to assess their possible disorders and as a result produce competent decision making, are crucial to ensure their health and ultimately enhance the level of public safety. In traditional intelligent health monitoring methods, the features are manually extracted depending on prior knowledge and diagnostic expertise. Inspired by the idea of unsupervised feature learning that uses artificial intelligence techniques to learn features from raw data, a two-stage learning method is proposed here for intelligent health monitoring of civil engineering structures. In the first stage, Nyström method is used for automatic feature extraction from structural vibration signals. In the second stage, Moving Kernel Principal Component Analysis (MKPCA) is employed to classify the health conditions based on the extracted features. In this paper, KPCA has been implemented in a new form as Moving KPCA for effectively segmenting large data and for determining the changes, as data are continuously collected. Numerical results revealed that the proposed health monitoring system has a satisfactory performance for detecting the damage scenarios of a three-story frame aluminum structure. Furthermore, the enhanced version of KPCA methods exhibited a significant improvement in sensitivity, accuracy, and effectiveness over conventional methods.

Keywords: damage detection; unsupervised feature learning; moving kernel principal component analysis; Nyström method

1. Introduction

Structural Health Monitoring (SHM) is defined as the process of gathering adequate information that allows detecting, locating and quantifying structural vulnerabilities early on fatigue cracking, degradation of boundary conditions, etc, thereby improving the resilience of the civil infrastructure (Boller *et al*. 2009).

Large-scale civil engineering structures, including buildings, bridges, dams, and pipelines, are lifelines for economic and social needs. Assessing their conditions and timely maintenance/ mitigation is critical for structural health monitoring to ensure their health, extend their service life, and ultimately enhance the level of public safety (Farrar and Worden 2012). Wireless sensor networks demonstrated great potential for full-spectrum SHM of the large-scale civil infrastructures (Alonso *et al*. 2018, Worden *et al*. 2007). These advanced sensor technologies are capable of collecting massive amounts of data.

Significant efforts on structural health monitoring have been extensively undertaken (Boller *et al*. 2009, Deraemaeker and Worden 2012, Farrar and Worden 2012). These studies could be mainly classified in two categories: (a) physics-based approaches which are based on vibratory characteristics of structural systems including natural frequency, mode and curvature (An *et al*. 2015, Ghiasi *et al*. 2016, 2017, Nobahari *et al*. 2017); and (b) data-driven approaches which extract sensitive features from sensor data to assess the structural conditions (Figueiredo *et al*. 2009, Krishnan *et al*. 2018, Malekzadeh *et al*. 2015, Wang and Cha 2017, 2018).

However, some of these studies may suffer three weaknesses as follows:

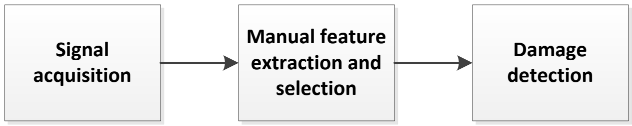
First, these approaches have not been implemented to remove the operational and environmental effects aggregated in extracted features; rather, they have been used to classify directly the extracted features in a supervised way, i.e., when data from the undamaged and damaged conditions are available. However, for most civil engineering infrastructures, where SHM systems are applied, the unsupervised learning algorithms are often required because only data from the undamaged condition are available (Cha and Wang 2017, Lei *et al*. 2016, Cha *et al*. 2017).

Second, the complexity and heterogeneity of the sensor data, post great challenges to data analysis for structural health monitoring and damage detection. Thus, it is in high demand in mining in-situ big data recorded from these large civil infrastructures and extracting sensitive features for damage detection (Deraemaeker and Worden 2012, Gui *et al*. 2017).

Third, traditional artificial intelligent techniques are unable to extract and organize the discriminative information from raw data directly. So lots of the actual efforts in intelligent diagnosis methods goes into the design

*Corresponding author, Professor,
E-mail: mrghasemi@eng.usb.ac.ir
[a] Ph.D. Student, E-mail: rghiasi@pgs.usb.ac.ir

http://www.techno-press.org/?journal=scs&subpage=6 ISSN: 1229-9367 (Print), 1598-6233 (Online)



(a) The traditional one

(b) A new one

Fig. 1 Intelligent health monitoring framework

of feature extraction algorithms in order to obtain the representative features from the signals. Such processes take advantage of human ingenuity but largely depend on much prior knowledge about signal processing techniques and diagnostic expertise, which is time-consuming and labor- intensive (Jia *et al*. 2016). Furthermore, lacking a comprehensive understanding of structural big data, it is often difficult to ensure the extracted features carrying optimal information to classify the mechanical faults. Thus, diagnosticians need to spend lots of time in analyzing these data and grasping their properties, being rather a tough task.

To overcome these weaknesses, a new two-stage framework for intelligent health monitoring of structure is presented here, as shown in Fig. 1(b). In this framework, the features are directly learned from structural raw signals and a classifier is used to classify the structural damage based on these learned features. The highlight of the framework is that the features are learned from the raw signals using a general-purpose learning procedure instead of being extracted by diagnosticians. The new framework releases the researcher from the tough tasks of designing feature extraction algorithms and it makes it easier to build a diagnosis system, therefore being more intelligent than the traditional one (Fig. 1(a)).

In the first stage of the proposed framework, Nyström method is used for automatic feature extraction from an array of accelerometers, when the structure operates in different structural state conditions (Yang *et al*. 2012). Then in the second stage, an unsupervised learning technique using moving kernel principal component analysis (MKPCA) is adapted for data normalization and damage detection. MKPCA is a revision of classical KPCA to make it more practical for long-term SHM. It models the effects of the operational and environmental variability on the extracted features. The algorithm produces a scalar output as a damage index, which should be nearly invariant when features are extracted from the normal condition. Finally, DIs from the feature vectors of the test data are classified through a threshold defined, based on the 95 percent cut-off value over the training data.

Fundamentally, PCA is a multivariable statistical method, based on the assumption of linearity in sets of variables. To some extent, many systems show a certain degree of nonlinearity and/or non-stationarity, and PCA may then overlook useful information on the nonlinear behavior of the system (Nguyen and Golinval 2010). As reported by Farrar *et al*. (2007), there are many types of damages that make an initially linear structural system respond in a nonlinear manner. Therefore, the detection problem necessitates methods that are able to study nonlinear systems, a reason for which motivate us to use KPCA.

KPCA is a nonlinear extension of PCA built to authorize features with nonlinear dependence among variables (Santos *et al*. 2016). Two main concerns inspired us to revise classical KPCA; one is the delay in damage detection and the other is the computational time issue (Cao *et al*. 2003). MKPCA, proposed in this paper, is to address and resolve these issues. It divides the large data into segments continuously collected and determines the sudden changes in windowed scheme.

Worden *et al*. (2007) have formulated seven axioms for SHM that capture general aspects that have emerged in several decades of experience. The following are two that are particularly relevant to this paper.

Axiom IVa: Sensors cannot measure damage. Feature extraction through signal processing and statistical classification is necessary to convert sensor data into damage information.

Axiom IVb: Without intelligent feature extraction, the more sensitive a measurement is to damage, the more sensitive it is to changing operational and environmental conditions.

The contributions of this paper are summarized as follows:

(1) Following the new intelligent health-monitoring framework, a two-stage learning method proposed in this paper. In the first learning stage, Nyström method is used to learn representative features from the structural vibration signals. Then in the second learning stage, MKPCA is trained to automatically classify the structural health conditions. Because of using an unsupervised scheme to learn the features, the proposed method does not depend on prior knowledge and human labor and may be more suitable for processing massive signals in the field of structural health monitoring.

(2) The performance of MKPCA is evaluated through receiver operating characteristic (ROC) curves. Its effectiveness is compared with several studies in the literature. ROC curves are used to determine performance on the basis of Type I/Type II error trade-offs. In SHM, in the context of damage detection, Types I and II errors are a false-positive and false-negative indication of damage, respectively (Santos *et al*. 2016).

(3) Another aim of the present work is to address the detection of nonlinear damage in civil structures using output-only measurements. The process of implementing nonlinear damage on a test bed structure will be explained in subsection 5.1. The detection is realized by comparing the reference and a current state of the system based on MKPCA output through the concept of subspace angle (Golub and Van Loan 1996).

The layout of the paper is as follows. First, mathematical description of the Nyström method will be explained. Next, the PCA method is reviewed briefly as it constitutes the starting point of the method. The kernel PCA method is described later to deal with nonlinear systems and the concept of subspace angle is introduced for damage detection purposes. Section 4 describes the flowchart of the proposed method. It combines the KPCA method with a windowed scheme in order to increase the sensitivity to damages. A description of the testbed structure, the simulated operational and environmental variability, and a summary of the data sets will be provided in Sub-Section 5.1. In 5.2, a comprehensive study is carried out using features extracted from time-series data sets measured with accelerometers deployed on the test bed structure. Finally, Section 6 highlights a discussion on the implementation and the analyses carried out in this paper, as a result of which some concluding remarks will be presented.

2. Automatic feature extraction by Nyström method

Automatic Feature Extraction (AFE) is a technique by which approximate eigenfunctions are deducted from gram matrices constructed via kernel functions. These eigenfunctions are features, engineered by a particular kernel (Golub and Van Loan 1996, Williams and Seeger 2001).

2.1 Mathematical background of Nyström method

Let $X_k \in \mathbb{R}^d, k = 1, \dots, n$ be a random sample drawn from a distribution $F(x)$ and $C \in \mathbb{R}^d$ be a compact set such that, $\mathcal{H} = \mathcal{L}^2(C)$ is a Hilbert space of functions given by the following inner product

$$\langle f, g \rangle_{\mathcal{H}} = \int f(x) g(x) dF(x) \tag{1}$$

Automatic Feature Extraction (AFE) using the Nyström method (Nyström 1930) aims at finding a finite dimensional approximation to the kernel eigenfunction expansion of Mercer kernels, as shown below.

$$K(x, t) = \sum_i \lambda_i \phi(x) \phi(t) \tag{2}$$

It is well known that Mercer kernels form a Reproducing Kernel Hilbert Space (RHKS) of functions (Rosipal and Trejo 2001). Every Mercer kernel defines a unique RHKS of functions as shown by the Moore-Aronszajn theorem (Aronszajn 1950). For a more involved treatment of RHKS and their applications, the reader may refer to the book written by Berlinet and Thomas-Agnan (2011).

Further, let $M(\mathcal{H}, \mathcal{H})$ be a class of linear operators from $\mathcal{H}$ to $\mathcal{H}$. Mercer's theorem (Mercer 1909) states that the spectral decomposition of the integral operator of K, $\mathcal{T} \in M(\mathcal{H}, \mathcal{H})$ defined below, yields the eigenfunctions which span the RHKS generated by K and having an inner product defined as Eq. (2).

$$(\mathcal{T}\phi_i)(t) = \int K(x, t) \phi(x) dF(x) \tag{3}$$

Eq. (3) is more common, known as the Fredholm integral equation of the first kind (Pérez-Rendón and Robles 2004). Nyström's method approximates this integral using the quadrature constructed by considering a finite kernel matrix constructed out of a prototype set $X_k, k = 1, \dots, m$. It calculates its spectral decomposition consisting of eigenvalues λ_k and eigenvectors u_k. This yields an expression for the approximate non-linear feature map $\hat{\phi}: \mathbb{R}^d \to \mathbb{R}^m$

$$\hat{\phi}_i(t) = \frac{\sqrt{m}}{\lambda_i} \sum_{k=1}^{m} K(X_k, t) u_{k,i} \tag{4}$$

3. Principal component analysis (PCA)

PCA is known as Karhunen-Loève transform or Proper Orthogonal Decomposition (POD) (Krzanowski 2000). For a dynamical system, a set of vibration features identified at a time t can be represented by the n-dimensional vector $x_k (k = 1, \dots, M)$, where M is the number of samplings and n is the number of output sensors. All the samples are collected in the observation matrix $X \in \Re^{n \times M}$. In general, PCA involves a data normalization procedure that leads to variables with zero-mean and unitary standard deviation.

PCA provides a linear mapping of data from the original dimension n to a lower dimension p using the transformation (Yan *et al*. 2005)

$$Y = TX \tag{5}$$

where $Y \in \Re^{p \times M}$ is called the score matrix and $T \in \Re^{p \times n}$ is the loading matrix. The dimension p represents the physical order of the system or the number of principal components affecting the vibration features. The loading matrix may be found from the main p eigenvectors of the covariance matrix X. In practice, PCA is often computed by Singular Value Decomposition (SVD) of the covariance matrix, i.e.

$$XX^T = U \sum U^T \tag{6}$$

where U is an orthonormal matrix whose columns define the principal components (PCs) and form a subspace spanning the data. Under certain assumptions, PCs in the matrix U may represent the vibration modes of the system (Feeny and Kappagantu 1998). As described in De Boe (2003) the order p of the system is determined by selecting the first p singular values in $\sum$ having a significant magnitude (energy). The effective number of PCs that is necessary for a good representation of the observation matrix X chosen based on of cumulated energies threshold. For practical purposes, a cumulated energy of 75% to 95%

is generally adequate for the selection of the active PCs (Nguyen and Golinval 2010).

3.1 Kernel Principal Component Analysis (KPCA)

The definitions and formulation presented here follow closely the ones described in De Boe (2003) and He *et al.* (2007). The reader may refer to these references for further details.

The key idea of KPCA is first to define a nonlinear map $x_k \phi(x_k)$with $x_k \in \Re^n, (K = 1, \ldots, M)$ which represents a high-dimensional feature space F, and then to apply PCA to the data in space F. With the assumption of centered data, i.e., $\sum_{i=1}^{M} \phi(x_i) = 0$, the covariance matrix in the space F is given by

$$C = \frac{1}{M} \sum_{i=1}^{M} \phi(x_i)\phi(x_i)^T \tag{7}$$

Principal components may be extracted next to solving the eigenvalue equation

$$\lambda V = CV \tag{8}$$

By defining the kernel matrix K of dimension $M \times M$ such that (Scholkopf *et al.* 1999)

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j) = \phi(x_i).\phi(x_j) \tag{9}$$

The eigenvalue problem may be put in the form

$$M\lambda\alpha = K\alpha \tag{10}$$

where α identifies the eigenvector V after normalization. The dynamical system characterized by the eigenvectors recognized in the feature space F and these eigenvectors can be considered as kernel principal components (KPCs). Note that, since the number of eigenvectors (i.e., nonlinear PCs) is the same as the number of samples, it is higher than the number of (linear) PCs given by PCA. The KPCA method is termed ‘‘nonlinear’’ since the feature mapping in the space F is achieved by a nonlinear function. The nonlinear property, extracted KPCs, should be able to reflect nonlinear or high order features to permit representation and classification of varied states. KPCA has the aptitude to use more nonlinear PCs to collect structural features than noise (Scholkopf *et al.* 1999). KPCA kernel can be chosen from a various function such as follows (Schölkopf *et al.* 1998):

- polynomial kernel,

$K(x_i, x_j) = (x_i.x_j + 1)^d$, where d is a positive integer;

- radial basis function (RBF),

$K(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / 2\sigma^2)$, where $2\sigma^2 = \gamma$ is the width of the Gaussian kernel.

In general, the above kernel functions give similar results if proper parameters are chosen. For instance, the width of the Gaussian kernel can be very small (< 1) or quite large. Therefore, the radial basis function present advantages owing to its flexibility in choosing the associated parameter. In contrast, the polynomial kernel requires a positive integer for the exponent.

3.2 Detection based on the concept of subspace angle

PCs define a subspace (or hyperplane) that characterizes the dynamic behavior of the system based on the active principal components. A change in the system modifies consequently its dynamic state and affects the subspace spanned by the PCs. This change may be estimated using the concept of subspace angle introduced by Golub and Van Loan (1996). Given two subspaces (each with linear independent columns) $S \in \Re^{n\times p}$ and $D \in \Re^{n\times q} (p > q)$, the procedure is as follows. Carry out the QR factorizations (Van Overschee and De Moor 2012)

$$S = Q_S R_S; D = Q_D R_D (Q_S \in \Re^{n\times p}, Q_D \in \Re^{n\times q}) \tag{11}$$

The columns of Q_S and Q_D define the orthonormal bases for S and D, respectively. The angles θ_i between the subspaces S and D are computed from singular values associated with the product $Q_S^T Q_D$

$$Q_S^T Q_D = U_{SD} \sum\nolimits_{SD} V_{SD}^T; \sum\nolimits_{SD} = diag\ (\cos(\theta_i)); i = 1, \ldots, q \tag{12}$$

The largest singular value is thus related to the largest angle characterizing the geometric difference between the two subspaces.

In SHM domain, damage may be detected by monitoring the angular coherence between subspaces estimated from the training and monitoring sets of the structure. A state is considered as a training set (reference state) if the system operates in normal conditions and damage does not exist in structure (i.e., nonlinearity is not activated). In addition, the monitoring set shows the current state of the system. Fig. 2 shows a 2D example in which an active subspace (or hyperplane) is built from two principal components.

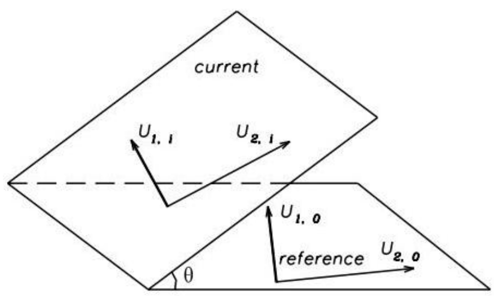


Fig. 2 Angle θ formed by active subspaces (hyper-planes) according to the training phase (reference state) and monitoring phase (current state), due to a dynamic change

4. The proposed framework for long-term monitoring

Real life employment of SHM involves dealing with a large amount of multivariate data. Only a small portion of abnormal data, in comparison to overall data, is available at the time when damage occurs. For detecting the changes in datasets effectively, the classical KPCA should be improved to make it more practical for long-term SHM data analysis. By means of KPCA, the damage can be detectable only when the principal components (eigenvectors) are influenced by abnormal behavior. Subsequently, eigenvectors are subject to changes only if a certain amount of abnormal data are captured and possibly affected the overall structure of data. This feature makes KPCA less effective for long-term SHM implementation. Therefore in this paper, Moving kernel principal component analysis (MKPCA) will be proposed to address this challenge.

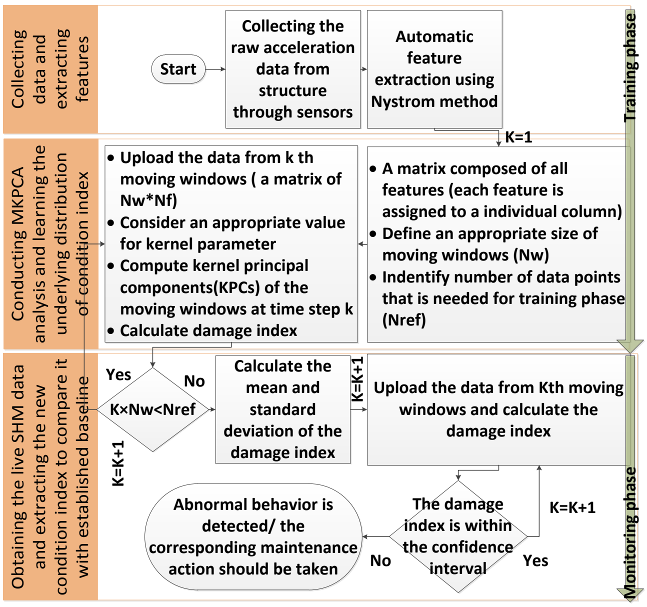


Fig. 3 Flowchart of the proposed method

4.1 Moving Kernel principal component analysis

Basically, MKPCA computes the KPCA within moving windows with a constant size. Fig. 3 gives details of MKPCA algorithm designed for long-term SHM applications. Moreover, for the sake of clarity, the entire process is summarized in Fig. 4. Also, the applied procedures can be summarized here as in three different steps as follows:

Step 1: A data matrix should be generated by sorting the time history data from each sensor into individual columns. Then, Nyström method will be used for an automatic feature extraction from raw acceleration data of each sensor. The feature matrix will be created based on the result of the Nyström method for all structural state conditions.

For each test, with consideration of N_s sensors and N_f extracted features from each sensor $N_s \times N_f$-dimensional feature vectors will be created when using them in the concatenated format.

Step 2: Feature matrix should be divided into two phases, training and monitoring. The training phase is intended for developing a baseline, a confidence interval, based on normal condition, while monitoring phase is set for a long-term monitoring. The dimensions of fixed moving windows should be well-defined. In fact, determining the window size precisely is one of the most critical issues in MKPCA, because it affects the speed and accuracy of the proposed method. In this paper, an efficient procedure will be implemented to choose the appropriate size of moving windows, which will be discussed in section 4.2 of this study. KPCA should be conducted for each window individually and results should be stored.

Step 3: A sensitive damage index should be selected in this step based on MKPCA outputs. In this paper, two different damage indexes are considered, the performance and robustness of which will be compared to various damage scenarios.

The first damage index is subspace angle defined earlier. In the second damage index, different entries of the first i principal components are considered as individual damage index is described as follow:

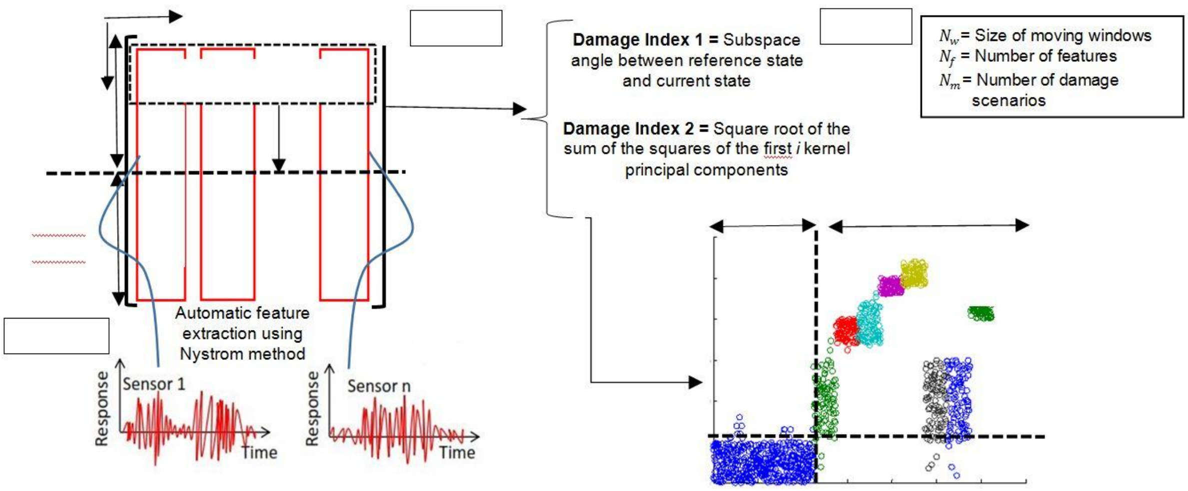


Fig. 4 A graphic representation of Moving Kernel Principal Component Analysis (MKPCA)

The Damage Indicator (DI_{rpc}) chosen for this study is the square root of the sum of the squares of the first i kernel principal components as shown in Eq. (13)

$$DI_{rpc} = \sqrt{(KPC_1)^2 + (KPC_2)^2 + \ldots + (KPC_i)^2} \quad i = 1, \ldots, n \tag{13}$$

where KPC_i is the i-th kernel principal components of moving windows and rpc is the abbreviation of the root of principal components. The reason for incorporating the first i kernel principal components in the DI_{rpc} is that they cover the most useful information in the data. In fact, since more than 99% of the energy distribution of KPCs is covered by the first i principal components, they are the only components incorporated in the DI_{rpc}.

It should be mentioned that the number of kernel principal components that should be considered depends on the data and there is not any prescription for all cases. However, in the most cases considered in this study, the most variance is covered by the first 5 components ($i = 5$) which will be discussed in more details in subsection 5.6. Therefore, if any damage occurs in the structure, it should affect the data and consequently variance of data and it should be detected by this damage index.

It should be noted that the DI_{rpc} for all windows are calculated w.r.t. time and thus their variations with time are plotted. As a final point, the confidence interval developed in the training phase should be considered as a benchmark (baseline) for detecting any possible damages in the rest of the data.

4.2 Size of the moving window and required data for training phase

The size of the window is a key parameter for the MKPCA algorithm. Commonly it is recommended that the window should be large enough so that the damage indexes are not influenced by the periodicity of the data. Alternatively, it should be small enough to timely detect abnormal behavior (Malekzadeh and Catbas 2016). This is a very general prescription and in most of the cases, there is not any straightforward procedure to identify the window-size. Therefore, in this section, a procedure is proposed as a method to define effective window-size for MKPCA.

Another important feature for an unsupervised damage detection algorithm is the amounts of data sets that are required for the training phase. The advantage is given to the algorithm that needs the least number of data sets to establish the baseline. The amount of data needed by an unsupervised algorithm for the training phase directly depends on the size of the window.

As a consequence, a yet new approach is proposed in this study to identify the appropriate size of the window and accordingly the required training data-sets. The sequential steps for this approach are presented as follow:

Step 1: The largest periodicity in the data (P) is identified.

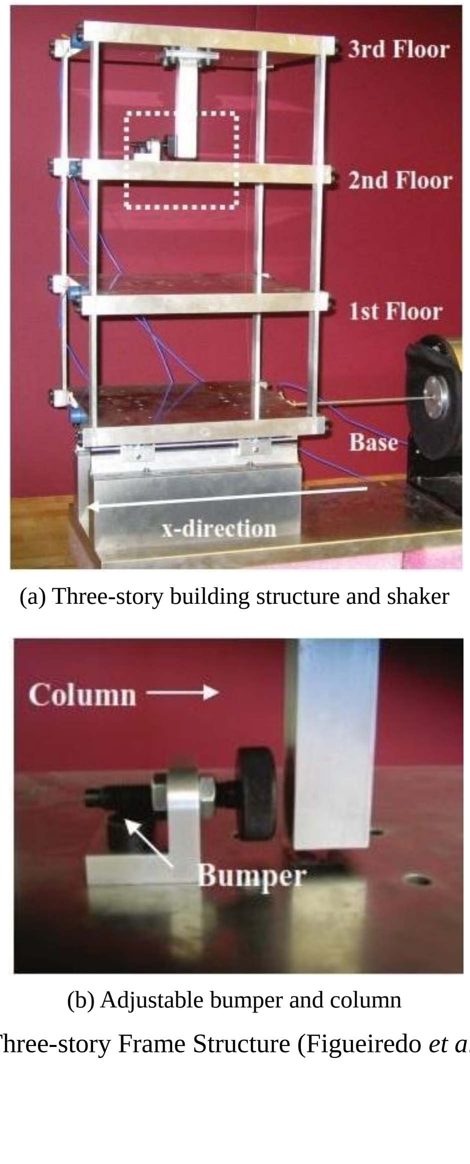
Step 2: A vector of window-size is generated as follows: Window Size = $[0.2P, 0.5P, \ldots, P, 2P]$

Step 3: The damage indexes are derived for all the window-sizes based on MKPCA. The optimized window-size is the smallest one that results in a stationary process.

5. Case study

5.1 Prototype structures and data

The standard data sets used in this study are from a three-story frame aluminum structure reported in Figueiredo *et al*. (2009) and has been intensively used for SHM validation in recent unsupervised damage detection approaches (Nguyen *et al*. 2014, Santos *et al*. 2016). Test bed building model is four-degree-of-freedom system with varied practical conditions, including variations in stiffness and mass loading. These variations simulate temperature changes and traffic, respectively. Those changes were designed to introduce variability in the fundamental natural frequency up to approximately 7 percent from the baseline condition, which is within the range normally observed in real-world structures (Figueiredo *et al*. 2011, Peeters and De Roeck 2001). Raw data were collected by four



(a) Three-story building structure and shaker

(b) Adjustable bumper and column

Fig. 5 Three-story Frame Structure (Figueiredo *et al*. 2009)

Table 1 Data labels of the structural state conditions

Label	Description
State 1	Baseline condition
State 2	Added mass (1.2 kg) at the base
State 3	Added mass (1.2 kg) on the 1st floor
State 4	States 4-9: 87.5% stiffness reduction at various positions to simulate temperature impact (more detail in Figueiredo *et al.*(2009))
State 5	
State 6	
State 7	
State 8	
State 9	
State 10	Gap (0.20 mm)
State 11	Gap (0.15 mm)
State 12	Gap (0.13 mm)
State 13	Gap (0.10 mm)
State 14	Gap (0.05 mm)
State 15	Gap (0.20 mm) and mass (1.2 kg) at the base
State 16	Gap (0.20 mm) and mass (1.2 kg) on the 1st floor
State 17	Gap (0.10 mm) and mass (1.2 kg) on the 1st floor

accelerometers mounted on the structure, as shown in Fig. 5.

A nonlinear damage scenario was introduced by contacting a suspended column with a bumper mounted on the floor below to simulate fatigue crack that can open and close under loading conditions or loose connections in structures. The smaller gap between the column and the bumper will result in the higher level of damage. Therefore, different levels of damage were created by adjusting the gap. More details about the test structure can be found in Figueiredo *et al*. (2009). Acceleration time-series from 17 different structural state conditions were collected, as described in Table 1, where the first 9 state conditions introduce the undamaged and the rest are damaged states. Time-series discretized into 4096 data points sampled at 3.125 ms intervals corresponding to a sampling frequency of 320 Hz.

For each structural state condition, data were acquired from 100 separate tests. Based on the test description (Figueiredo and Flynn 2009), state1 is the baseline condition (reference state) of the structure and states 2-9 include those states with simulated operational and environmental variability. State14 is considered as the most severely damaged one as it corresponds to the smallest gap case, which induces the highest number of impacts. State10 is the least severe damaged scenario and states11-13 represent mid-level damage scenarios. States15-17 are the variant states of either state10 or state13 with mass added effect in order to create more realistic conditions.

It is worth noting that, although the developed data-driven machine learning methods presented in Section 4 were herein verified using a test bed found in the literature, it is capable to detect damage in other civil, infrastructure and mechanical systems with least modifications, which is the best feature in data-driven methods over conventional physical-based methods.

5.2 Experimental results and analysis

In this study, the features extracted by the Nyström method from response time series are used as damage-sensitive features. Thus, for each test at each state condition, the features are extracted using time-series from all four accelerometers (channels 2-5) and stored in a feature vector. For each test, the number of extracted parameters from each sensor is 10, therefore when using them in concatenated format, it yields to 40-dimensional feature vectors. This analysis is based on the assumption of an output-only damage detection approach, and so data from the channel 1 (the input force) is not used.

Note that extracted parameters should be constant when estimated based on time-series data obtained from time-invariant systems. However, in the presence of operational and environmental conditions as well as damage, the parameters are expected to change accordantly. To better understand such changes, the next section will investigate it in more details.

5.3 Sensitivity and amplitude of features

To better understand the performance of the proposed machine learning techniques, selection of effective and sensitive damage features are in high demand, since they could theoretically perform sensitive and robust media to all kinds of damages even under high variations and other interferences. The 3rd floor’s sensor data have been selected as the benchmark to demonstrate our concept.

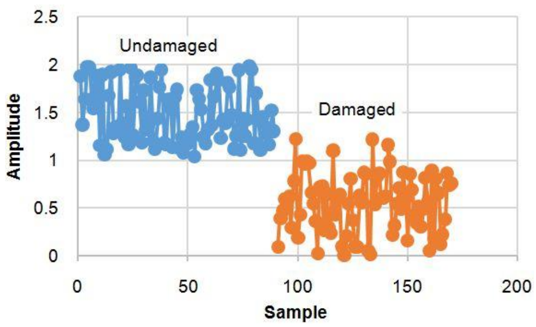
Figs. 6(a) and (b) are plotted for the first and second features, extracted automatically using Nyström method to demonstrate the level/amplitude of different feature parameters for undamaged or damage states. The Nyström method is configured to use the RBF kernel with parameter $\gamma = 0.75$ (De Brabanter *et al*. 2010, Chandorkar *et al*. 2015). These plots are composed of an extracted feature from 10 out of 100 tests of each undamaged state (states1-9) and 10 out of 100 tests of each damaged state (states10-17), so it has a dimension of 170_1.

The Extracted Feature 1 (EF1) and EF2 are found as the best sensitive features to distinguish between undamaged and damaged states. As clearly illustrated in Figs. 6(a) and (b), two plots have different thresholds but both exhibit significant differences in data trend to allow clear identification of undamaged or damaged structures.

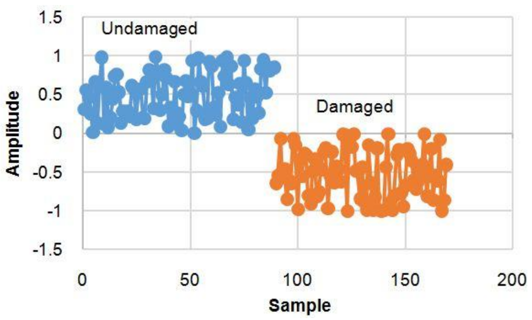
5.4 Comparing damage indexes

There are several parameters in the proposed method, i.e., the selected damage index, the optimized size of moving windows, and a number of considered KPCs. A study of parameters selection will be carried out here accordingly.

First, the effect of various damage indexes on the accuracy of the method will be investigated. As stated in section 4.1, two damage sensitive indexes, DI_{rpc} and subspace angle is used. Training data set is composed of an extracted feature from 10 out of 100 tests of each



(a) EF_1



(b) EF_2

Fig. 6 Amplitudes from damaged and undamaged condition

undamaged state (states1-9) and testing data set is composed of feature from 10 out of 100 tests of each undamaged and damaged state (states1-17). It should be noted that 20 trials are carried out for each experiment in the following study to reduce the effects of the randomness. The damage detection results are shown in Fig. 7. In the Fig., the training accuracies, testing accuracies, training time and the testing time are averaged by 20 trials, and the positive error bars show the standard deviations. The computation platform is a PC with an Intel Core™ i7 2.0 GHz and 4G RAM.

It can be seen that all training accuracies are over 93.0% and testing accuracies are over 90.08%. It means that the proposed method is able to classify the 9 damage conditions of the case study dataset with high accuracies using various damage indexes.

Moreover, the standard deviations of training accuracies and testing accuracies using the subspace angle method are small, indicating that the damage conditions can be detected reliably. However, comparing to DI_{rpc}, subspace angle spends more time. On the contrary, the standard deviations of training and testing accuracies using the DI_{rpc} are larger than subspace angle and as a result, the reliability of the method will be reduced.

5.5 Effect of appropriate size selection of moving windows

After evaluating the selection of damage index on the accuracy of the proposed method, the effect of a proper size selection of moving windows on the accuracy of MKPCA will be assessed in this sub-section. A total of 17 data sets were considered in this study. Given that for each structural state condition 100 separate tests considered, the maximum periodicity of data points (P) is 100. This results in the main matrix with 1700 rows as the feature vectors and 40 columns as the extracted features. The corresponding results for the MKPCA are visualized in Fig. 8. The horizontal axis indicates state condition of structure while the vertical axis represents the F-score for three sizes of the moving window.

F-score is a commonly used criterion measuring the performance of a classification method (Sokolova *et al*. 2006). It considers both the precision (p) and the recall (r) of the result to compute the score. p is the number of correct positive results divided by the number of all positive results returned by the classifier, and r is the number of correct positive results divided by the number of all true positives. In other words, the precision is a function of the correctly classified samples, denoted as true positives, TP, and samples misclassified as positives (false positives, FP), and recall is a function of TP and its misclassified samples (false negatives, FN). F-score reaches its best value at 1 and the worst score is at 0.

$$p = \frac{TP}{TP + FP} \; and \; r = \frac{TP}{TP + FN} \tag{14}$$

$$F - score = 2 \times \frac{p \times r}{p + r} \tag{15}$$

In Fig. 8, the F-scores of the case study dataset using the three size of moving windows are shown. It can be seen that the F-scores using N_w = 10 range from 0.953 to 1, whereas

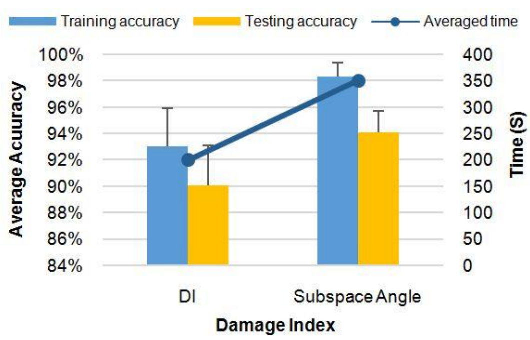


Fig. 7 Damage detection results using various damage indexes

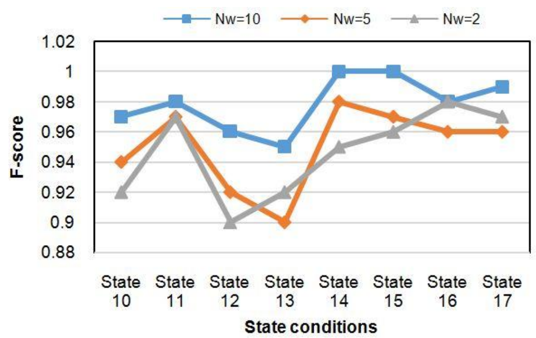


Fig. 8 F-score of data set for selecting optimized size of moving windows

Table 2 Number and percentage of Type I for each size of moving windows

Windows size based on periodicity (P)	Size of moving windows (N_w)	Type I error
0.02P	2	32 (7.11%)
0.05P	5	20 (4.44%)
0.1P	10	16 (3.56%)
0.5P	50	28 (6.22%)
P	100	46 (10.22%)
2P	200	70 (15.56%)

the F-scores using N_w = 5 range from 0.907 to 0.98 and the F-scores using the method N_w = 2 range from 0.899 to 0.98.

The appropriate size of moving windows can also be selected based on Type I error that shows false indication of damage on date set of undamaged state. Table 2 shows results when training matrix is composed of extracted parameters from 50 out of 100 tests from each undamaged state (states1-9), and the test matrix Z is composed of the remaining 50 tests of each undamaged state.

For most health conditions, the N_w = 10 obtains higher F-scores than the other size. Furthermore, N_w = 10 has smallest Type I error. Taking this information into account, the size of the moving window is chosen as 10×40. Hence, it means 10 feature vectors must be used for processing and interpreting of the data.

5.6 KPCA-based detection method

KPCA differs from PCA, notably in the number of kernel principal components (KPCs) which is equal to the number M of time samples whereas the number of PCs in the PCA method is equal to the number n of measurement responses (Cao *et al*. 2003). In the KPCA method, the energy distribution of KPCs depends on the parameter chosen in the kernel function. Furthermore as discussed in section 4, most of the energy distribution of KPCs is covered by the first principal components. Therefore in this sub-section, results were achieved with KPCA using the RBF kernel with a parameter $\gamma = 0.025$. The eigenvalue diagram of Fig. 9 shows that the first five KPCs apprehend most of the system energy.

It must be noted while PCs represent modal features (e.g., mode-shapes) under certain circumstances, KPCs do

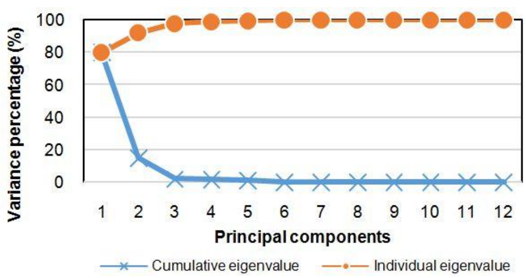


Fig. 9 Eigenvalue diagram

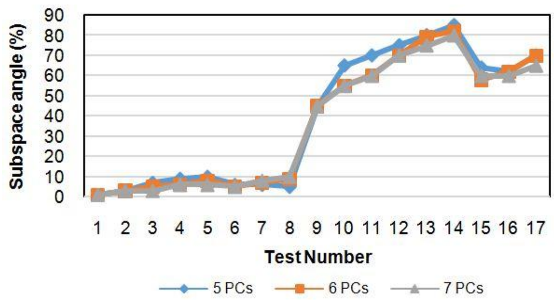


Fig. 10 MKPCA detection based on the subspace angle

not represent any specific physical meanings (He *et al*. 2007). However, KPCs are useful for classifying different dynamic behaviors of the system. The KPCA method needs to add up a sufficiently large number of KPCs in the subspace to accumulate enough effective information. On the other hand, KPCA offers a possibility of redistributing the energy of KPCs i.e. to regulate the main energy on the first but also on secondary KPCs (Nguyen and Golinval 2010). Noted that noise is related mainly to last KPCs. Therefore, KPCA is more immune to noise because of its capability to extract several first representative KPCs.

To compare the performance of MKPCA using the various number of KPCs, one test from each structural state condition is considered. As shown in Fig. 10, the use of five KPCs allows a good detection of various damage state. The result is based on subspace angle index. It also shows that, by taking into account some more KPCs, the detection is actually performed in a stable manner. In this example, the optimum detection is achieved with 5 KPCs. In conclusion, the KPCA-based method is effective for the detection of the nonlinear damage.

5.7 Assessing the proposed approach in term of feature extraction

To evaluate the effectiveness of the proposed methods for AFE, several existing studies in the literature are used for comparison, as listed in Table 3 (Gui *et al*. 2017, Santos *et al*. 2016). Clearly, different methods exhibit somehow different levels of accuracy, ranging up to five percent deviation for feature extraction and damage detection prediction. For example, the Auto-associative neural network has error of 4.3 percent in prediction and the one-class SVM reaches up to 3.36 percent deviation to true solutions. As a comparison, the SVM with optimization technique (GS) could yield a better prediction when using auto-regressive (AR) features within 1.18 percent. It may confirm that the use of optimization techniques could help to select the parameters for damage detection. Meanwhile as stated in previous sections traditional artificial intelligent techniques are unable to extract and organize the discriminative information from raw data directly. Thus, some actual efforts in intelligent damage detection methods are focused on the design of feature extraction algorithms in order to obtain the representative features from the signals (Krishnan *et al*. 2018, Langone *et al*. 2017). To overcome this problem, AFE based on Nyström method is used, which further increases accuracy as shown in Table 3. It can be

Table 3 Comparison of the proposed study with existing ones in the literature

	Method	Feature extraction method	Prediction error
Exiting studies	Auto-associative neural network (AANN)*	AR Parameter	4.30 %
	Factor analysis (FA)*		4.20 %
	Mahalanobis squared distance (MSD)*		4.00 %
	Singular value decomposition (SVD)*		4.60 %
	One-class SVM**		3.36 %
	Support vector data description (SVDD)**		3.44 %
	Kernel principal component analysis (KPCA)**		2.72 %
	Greedy kernel principal component analysis (GKPCA)**		2.64 %
	Particle Swarm Optimization (PSO)+ SVM***		2.35 %
	Genetic Algorithm (GA)+ SVM***		2.35 %
	Grid-search techniques (GS)+ SVM***		1.18 %
Proposed study	Nyström method+ MKPCA	Automatic Feature Extraction (AFE)	0.42 %

envisioned that the proposed machine learning with AFE will be more robust when under those complex data. A comprehensive comparison will be discussed in the next section.

5.8 Accuracy of damage detection methods

Having chosen the parameters of the proposed method, a comprehensive comparison will be made here between different methods to evaluate the effectiveness of the MKPCA method.

For generalization purposes, the feature vectors are split into the training and test matrices. The training matrix, X, permits each algorithm to learn the underlying distribution and dependency of all undamaged states on the simulated operational and environmental variability. Thus, this matrix is composed of extracted parameters from 50 out of 100 tests from each undamaged state (states1-9), and so it has a dimension of 450 _40. The test matrix Z (1250 _40) is composed of extracted parameters from the remaining 50 tests of each undamaged state together with extracted parameters from all the 100 tests of each damaged state (states10-17). This procedure permits one to evaluate the generalization performance of the machine learning algorithms in an exclusive manner because time-series used in the test phase are not included in the training phase. During the test phase, the algorithms are expected to detect deviations from the normal condition when feature vectors come from damaged states, even in the presence of operational and environmental effects.

The next step is to carry out statistical modeling for feature classification. In that regard, the algorithm based on MKPCA is implemented in an unsupervised learning mode by first taking into account features from all the undamaged state conditions (training matrix). To evaluate the effectiveness of the proposed methods several existing studies in the literature (Gui *et al*. 2017, Nguyen *et al*. 2014, Santos *et al*. 2016) are used for comparison. These existing methods could be classified in two categories: (a) kernel-based algorithms such as one-class SVM, KPCA, GKPCA; and (b) enhanced versions of SVM with optimization algorithm (Gui *et al*. 2017): SVM-PSO, SVM-GA, and SVM-GS, which implemented and configured as described in Nguyen *et al*. (2014) and Santos *et al*. (2016). Finally, for each algorithm, the damage index is stored into a 1250-length vector.

The subset of 25 percent of the training data is used for MKPCA kernel projection and algorithm is configured to retain 90 percent of the variability in the data after dimension reduction. Furthermore, the RBF kernel with parameter $\gamma = 0.025$ is chosen for KPCA kernel (Gui *et al*. 2017).

ROC curves are used to compare the performance of various methods based on the tradeoff between Type I and Type II error. The point at the left-upper corner of the plot (0, 1) is called a perfect classification. Figs. 11 and 12 illustrate the ROC curves for aforementioned algorithms.

In the first case, all kernel-based algorithms are compared together (Fig. 11). The curves show that none of the algorithms can have a perfect classification with a linear threshold because none of the curves goes through the left-upper corner, neither have supremacy in terms of true detection rate for the entire false alarm domain. Furthermore, for levels of significance around 5 percent, the GKPCA and MKPCA have better true detection rate than one-class SVM and KPCA. They are in fact the ones that maximize the true detection of damaged cases with similar performances in terms of false alarm rate. It worth to be noted that, false alarm rate of 0.05 is acceptable in real-world scenarios of SHM (Santos *et al*. 2016).

Note that the proposed algorithms apply data transformation in the high dimensional feature space to achieve a data model that represents the normal structural condition. On the other hand, the MKPCA algorithm reduces the data dimensionality in high-dimensional space, storing the principal components that have the largest variability in the data.

Due to the superiority of the MKPCA to the KPCA, the authors chose the MKPCA algorithm for a comparative analysis with improved algorithms already made known in the literature (Gui *et al*. 2017). In the second case, Fig. 12 shows a comparison between the SVM+PSO, SVM+GS, SVM+ GA and MKPCA algorithms. In the whole false alarm range and for a given threshold, the proposed

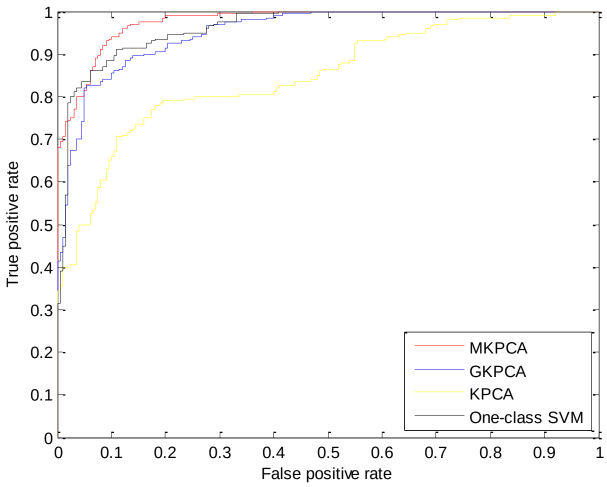


Fig. 11 ROC curves for the GKPCA, KPCA, one-class SVM and MKPCA algorithms

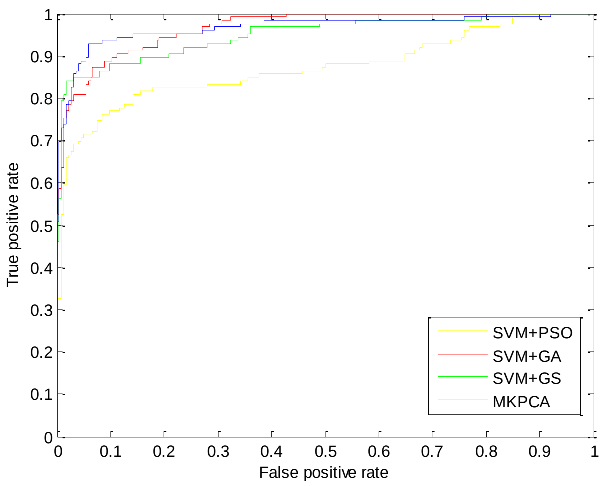
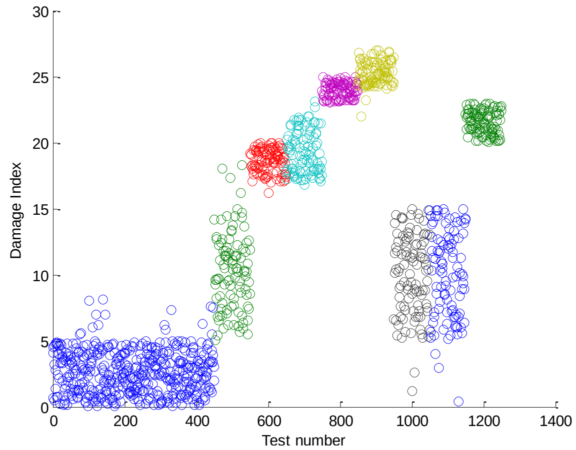
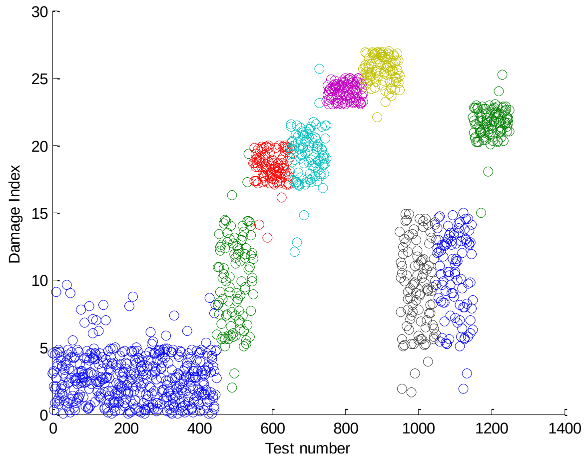


Fig. 12 ROC curves for the SVM+PSO, SVM+GA, SVM+GS and MKPCA algorithms



(a) MKPCA



(b) Original KPCA

Fig. 13 DI_{rpc} calculated based on feature vectors from the undamaged and damaged condition using two version of KPCA algorithms along with defined threshold

algorithm has, clearly, better performance to detect abnormal conditions in the test structure than the improved SVM-based algorithms. In general, the MKPCA seems to be more effective than the SVM in most points of the ROC curves by assimilating the normal structural condition embedded into the principal components, especially for false alarm rates around 5 percent.

In order to quantify the performance of the MKPCA and KPCA for a given threshold, Figs. 13(a) and (b) plot the DI_{rpc} for the feature vectors of the entire test data along with a threshold defined based on the 95 percent cut-off value over the training data. Both algorithms show a monotonic relationship between the level of damage and amplitude of the DI_{rpc}, even when operational and environmental variabilities are present. In the other words, the approaches are able to remove the operational and environmental effects in such a way that DI_{rpc} from states15, 16 and 10 have similar amplitude, as well as state17 is associated with state13. As stated in subsection 5.1, states15-17 are the variant states of either state10 or state13 with operational effects.

The performance of statistical classification methods mostly indicated by the Type I (a false-positive indication of damage) and Type II (false-negative indication of damage) errors. This technique recognizes that a false-positive classification may have different consequences than false-negative one (Santos *et al*. 2016). In Figs. 13(a) and (b), the Type I errors are DI_{rpc} that exceed the threshold value in the undamaged condition domain (1-450). On the other hand, the Type II errors are DI_{rpc} that does not surpass the threshold value in the damaged condition domain (451-1250). Table 4 summarizes the number and percentage of Type I and Type II errors for all algorithms that mentioned in this paper.

In terms of an overall analysis, the proposed scheme based on AFE and MKPCA has the best overall performance in detecting damage (0.625%) and representing the normal condition (3.55%). On the contrary, The FA algorithm has a good performance to avoid false indications of damage (2.22 %) but has the worst performance to detect damage (5.38%), due to its being sensitive to the number of factors driving changes in

Table 4 Number and percentage of Type I and Type II errors for each algorithm

Algorithm	Error			Accuracy	F1-score
	Type I	Type II	Total		
ANN*	44 (9.78%)	10 (1.25%)	54 (4.32%)	0.956	0.966
FA*	10 (2.22%)	43 (5.38%)	53 (4.24%)	0.957	0.966
MSD*	42 (9.33%)	8 (1.00%)	50 (4.00%)	0.96	0.969
SVD*	29 (6.44%)	29 (3.62%)	58 (4.64%)	0.953	0.963
One-class SVM*	36 (8.00%)	6 (0.75%)	42 (3.36%)	0.966	0.974
SVDD*	37 (8.22%)	6 (0.75%)	43 (3.44%)	0.965	0.973
KPCA*	21 (4.67%)	13 (1.62%)	34 (2.72%)	0.972	0.978
GKPCA*	24 (5.33%)	9 (1.12%)	33 (2.64%)	0.973	0.979
AFE+ MKPCA	16 (3.55%)	5 (0.625%)	21 (1.68%)	0.983	0.987

features.

The one-class SVM and SVDD algorithms performed similarly and offered 0.75% performance to detect damage; however, their false alarm rate is relatively high as greater than 8%. This can be due to that, the selected support vectors can easily identify the damaged data, but are not properly representing the normal condition as the number of Type I errors is higher than 5 percent.

In the field of damage detection, sensitivity measures the portion of damaged cases correctly identified and specificity measures the portion of undamaged cases which are correctly identified. According to this definition, the AANN reveals the worst performance in terms of Type I errors (9.78 percent), but a relatively good performance in terms of minimization of Type II errors (1.25 percent), which indicates that when the sensitivity of the classifier is increased, and so it detects more damaged scenarios, it also increases the number to mislabels of undamaged cases. Thus, as the sensitivity goes up, specificity goes down.

The KPCA and GKPCA attempt a balancing between Type I and Type II errors, their total error is relatively small, which is supported by the retention of principal components in the high-dimensional space. This should eliminate variability caused by operational and environmental effects. Finally, the proposed algorithm has a tendency to reduce the total number of misclassifications (1.68 percent) when compared with other aforementioned algorithms. This superiority might be related to the ability of the proposed algorithm to find nonlinear patterns in the data via the kernel trick, as well as the independence on the choice of the initial parameters.

Nevertheless, as far as the total misclassifications is concerned, all algorithms perform relatively well on these standard data sets for Type I and Type II errors. Additionally, based on Table 4 and for these specific data sets, Two important conclusions can be drawn: (a) when life safety issues are the main reason for deploying an SHM system and one wants to minimize false-negative indications of damage, the proposed scheme, one-class SVM and SVDD algorithms are preferred. (b) When reliability issues are ruling an SHM system and one wants to minimize false-positive indications of damage without increasing the false-negative indications of damage, the proposed scheme, the GKPCA, and KPCA algorithms are more appropriate. Finally, MKPCA has good generalization performance, which is a very important advantage for real-world applications, where the thresholds are defined based on undamaged data used in the training phase.

6. Conclusions

In this paper, the performance of the proposed two-stage method for structural damage detection, under varying operational and environmental conditions, is shown using benchmark data sets from a well-known base-excited three-story frame structure. The data sets are characterized by 17 different structural state conditions, including linear changes caused by varying stiffness and mass-loading conditions as well as nonlinear effects caused by damage. Different levels of damage were created by adjusting the gap between a suspended column and a bumper.

It would be desirable to make an intelligent health monitoring less dependent on prior knowledge and diagnostic expertise when processing structural big data, especially in the feature extraction step. In this paper, the Nyström method adaptively studies features that capture discriminative information from vibration signals in an unsupervised way. Then the features are fed to MKPCA to classify health conditions. Through the case studies on the frame structure, it is shown that the proposed method adaptively learns features from raw signals for various diagnosis issues and is superior to the existing methods in health monitoring of the structure.

The MKPCA algorithm is shown to be reliable to create a global damage index that can separate damaged from undamaged conditions, even when the structure is operating under varying operational and environmental conditions and yield better results in terms of minimization of total misclassifications; three fundamental reasons are also proposed for that behavior:

(i) All kernel-based algorithms map the original observations into the high-dimensional space. However, the MKPCA map the original observations into the high-dimensional space, in order to model nonlinear patterns presented in the original observation space and capture the operational and environmental effects with a known percentage. They are then back to the original space to perform the damage detection.
(ii) The KPCA perform the damage detection by retaining the principal components that take into account 90 percent of the data variability. It might be useful to discard some sort of noise and singularities from the data that can mask changes caused by damage from changes caused by operational and environmental conditions. The KPCA eliminate some noise from the data and provide a better trade-off between sensitivity and specificity.
(iii) In order to make KPCA method more practical and

feasible for long-term monitoring, a windowing technique is employed. As a result, MKPCA algorithm proposed as promising upgraded version of KPCA that can segment the data flow to detect damage more precisely.

Finally, MKPCA algorithms do not require a direct measure of the sources of variability, e.g., traffic loading and temperature. Instead, the algorithm relies only on measured response time-series data acquired under varying operational and environmental effects.

References

Alonso, L., Barbarán, J., Chen, J., Díaz, M., Llopis, L. and Rubio, B. (2018), "Middleware and communication technologies for structural health monitoring of critical infrastructures: A survey", *Comput. Stand. Interf.*, **56**(Supplement C), 83-100.

An, D., Kim, N.H. and Choi, J. (2015), "Practical options for selecting data-driven or physics-based prognostics algorithms with reviews", *Reliab. Eng. Syst. Saf.*, **133**, 223-236.

Aronszajn, N. (1950), "Theory of reproducing kernels", *Trans. Am. Math. Soc.*, **68**(3), 337-404.

Berlinet, A. and Thomas-Agnan, C. (2011), *Reproducing kernel Hilbert spaces in probability and statistics*, Springer Science & Business Media.

Boller, C., Chang, F.-K. and Fujino, Y. (2009), *Encyclopedia of structural health monitoring*, John Wiley & Sons.

Cao, L.J., Chua, K.S., Chong, W.K., Lee, H.P. and Gu, Q.M. (2003), "A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine", *Neurocomputing*, **55**, 321-336.

Cha, Y.-J. and Wang, Z. (2017), "Unsupervised novelty detection-based structural damage localization using a density peaks-based fast clustering algorithm", *Struct. Heal. Monit.*, **17**(2), 313-324.

Cha, Y., Choi, W. and Büyüköztürk, O. (2017), "Deep learning-based crack damage detection using convolutional neural networks", *Comput. Civ. Infrastruct. Eng.*, **32**(5), 361-378.

Chandorkar, M., Mall, R., Lauwers, O., Suykens, J.A.K. and De Moor, B. (2015), "Fixed-size least squares support vector machines: Scala implementation for large scale classification", *2015 IEEE Symposium Series on Computational Intelligence*, pp. 522-528.

De Boe, P. (2003), "Les éléments piézo-laminés appliqués à la dynamique des structures", Ph.D. Dissertation.

De Brabanter, K., Karsmakers, P., Ojeda, F., Alzate, C., De Brabanter, J., Pelckmans, K., … Suykens, J.A.K. (2010), LS-SVMlab Toolbox User's Guide: version 1.7. Katholieke Universiteit Leuven.

Deraemaeker, A. and Worden, K. (2012), *New trends in vibration based structural health monitoring* (Vol. 520), Springer Science & Business Media.

Farrar, C.R. and Worden, K. (2012), *Structural health monitoring: a machine learning perspective*, John Wiley & Sons.

Farrar, C.R., Worden, K., Todd, M.D., Park, G., Nichols, J., Adams, D.E., Bement, M.T. and Farinholt, K. (2007), *Nonlinear system identification for damage detection*, Los Alamos National Laboratory (LANL), Los Alamos, NM, USA.

Feeny, B.F. and Kappagantu, R. (1998), "On the physical interpretation of proper orthogonal modes in vibrations", *J. Sound Vib.*, **211**(4), 607-616.

Figueiredo, E. and Flynn, E. (2009), "Three-story building structure to detect nonlinear effects", *Rep. SHMTools Data Descr.*

Figueiredo, E., Park, G., Figueiras, J., Farrar, C. and Worden, K. (2009), *Structural health monitoring algorithm comparisons using standard data sets*, Los Alamos National Laboratory (LANL), Los Alamos, NM, USA.

Figueiredo, E., Park, G., Farrar, C.R., Worden, K. and Figueiras, J. (2011), "Machine learning algorithms for damage detection under operational and environmental variability", *Struct. Heal. Monit.*, **10**(6), 559-572.

Ghiasi, R., Torkzadeh, P. and Noori, M. (2016), "A machine-learning approach for structural damage detection using least square support vector machine based on a new combinational kernel function", *Struct. Heal. Monit.*, **15**(3), 302-316.

Ghiasi, R., Ghasemi, M.R. and Sohrabi, M.R. (2017), "Structural Damage Detection using Frequency Response Function Index and Surrogate Model Based on Optimized Extreme Learning Machine Algorithm", *J. Comput. Methods Eng.*, **36**(1), 1-17.

Golub, G.H. and Van Loan, C.F. (1996), *Matrix computations. 1996*, Johns Hopkins University Press, Balt, MD, USA, pp. 374-426.

Gui, G., Pan, H., Lin, Z., Li, Y. and Yuan, Z. (2017), "Data-driven support vector machine with optimization techniques for structural health monitoring and damage detection", *KSCE J. Civ. Eng.*, **21**(2), 523-534.

He, Q., Kong, F. and Yan, R. (2007), "Subspace-based gearbox condition monitoring by kernel principal component analysis", *Mech. Syst. Signal Process.*, **21**(4), 1755-1772.

Jia, F., Lei, Y., Lin, J., Zhou, X. and Lu, N. (2016), "Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data", *Mech. Syst. Signal Process.*, **72**, 303-315.

Krishnan, M., Bhowmik, B., Hazra, B. and Pakrashi, V. (2018), "Real time damage detection using recursive principal components and time varying auto-regressive modeling", *Mech. Syst. Signal Process.*, **101**, 549-574.

Krzanowski, W. (2000), *Principles of multivariate analysis— a user's perspective*, Oxford University Press.

Langone, R., Reynders, E., Mehrkanoon, S. and Suykens, J.A.K. (2017), "Automated structural health monitoring based on adaptive kernel spectral clustering", *Mech. Syst. Signal Process.*, **90**, 64-78.

Lei, Y., Jia, F., Lin, J., Xing, S. and Ding, S.X. (2016), "An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data", *IEEE Trans. Ind. Electron.*, **63**(5), 3137-3147.

Malekzadeh, M. and Catbas, F.N. (2016), "A Machine Learning Framework for Automated Functionality Monitoring of Movable Bridges", In: *Dynamics of Civil Structures*, **2**, 57-63.

Malekzadeh, M., Atia, G. and Catbas, F.N. (2015), "Performance-based structural health monitoring through an innovative hybrid data interpretation framework", *J. Civ. Struct. Heal. Monit.*, **5**(3), 287-305.

Mercer, J. (1909), "Functions of positive and negative type, and their connection with the theory of integral equations", *Philos. Trans. R. Soc. London. Ser. A, Contain. Pap. a Math. or Phys. Character*, **209**, 415-446.

Nguyen, V.H. and Golinval, J. (2010), "Fault detection based on Kernel Principal Component Analysis", *Eng. Struct.*, **32**(11), 3683-3691.

Nguyen, T., Chan, T.H.T. and Thambiratnam, D.P. (2014), "Controlled Monte Carlo data generation for statistical damage identification employing Mahalanobis squared distance", *Struct. Heal. Monit.*, **13**(4), 461-472.

Nobahari, M., Ghasemi, M.R. and Shabakhty, N. (2017), "Truss structure damage identification using residual force vector and genetic algorithm", *Steel Compos. Struct., Int. J.*, **25**(4), 485-496.

Nyström, E.J. (1930), "Über die praktische Auflösung von Integralgleichungen mit Anwendungen auf Randwertaufgaben", *Acta Math*., **54**(1), 185-204.
Peeters, B. and De Roeck, G. (2001), "One-year monitoring of the Z 24-Bridge: environmental effects versus damage events", *Earthq. Eng. Struct. Dyn*., **30**(2), 149-171.
Pérez-Rendón, A.F. and Robles, R. (2004), "The convolution theorem for the continuous wavelet tranform", *Signal Processing*, **84**(1), 55-67.
Rosipal, R. and Trejo, L.J. (2001), "Kernel partial least squares regression in reproducing kernel hilbert space", *J. Mach. Learn. Res*., **2**(Dec), 97-123.
Santos, A., Figueiredo, E., Silva, M.F.M., Sales, C.S. and Costa, J.C.W.A.C.W.A. (2016), "Machine learning algorithms for damage detection: Kernel-based approaches", *J. Sound Vib*., **363**, 584-599.
Schölkopf, B., Smola, A. and Müller, K.-R. (1998), "Nonlinear component analysis as a kernel eigenvalue problem", *Neural Comput*., **10**(5), 1299-1319.
Scholkopf, B., Mika, S., Burges, C.J.C., Knirsch, P., Muller, K.-R., Ratsch, G. and Smola, A.J. (1999), "Input space versus feature space in kernel-based methods", *IEEE Trans. Neural Networks*, **10**(5), 1000-1017.
Sokolova, M., Japkowicz, N. and Szpakowicz, S. (2006), "Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation", In: *Australian Conference on Artificial Intelligence*, pp. 1015-1021.
Van Overschee, P. and De Moor, B.L. (2012), *Subspace identification for linear systems: Theory—Implementation—Applications*, Springer Science & Business Media.
Wang, Z. and Cha, Y.-J. (2017), "Unsupervised Novelty Detection Techniques for Structural Damage Localization: A Comparative Study", In: *Model Validation and Uncertainty Quantification*, **3**, 125-132.
Wang, Z. and Cha, Y. (2018), "Automated damage-sensitive feature extraction using unsupervised convolutional neural networks", In: *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2018*, **10598**, 105981J.
Williams, C.K.I. and Seeger, M. (2001), "Using the Nyström method to speed up kernel machines", In: *Advances in Neural Information Processing Systems*, pp. 682-688.
Worden, K., Farrar, C.R., Manson, G. and Park, G. (2007), "The fundamental axioms of structural health monitoring", *Proc. R. Soc. London A Math. Phys. Eng. Sci*., **463**(2082), 1639-1664.
Yan, A.-M., Kerschen, G., De Boe, P. and Golinval, J.-C. (2005), "Structural damage diagnosis under varying environmental conditions—part I: a linear analysis", *Mech. Syst. Signal Process*., **19**(4), 847-864.
Yang, T., Li, Y.-F., Mahdavi, M., Jin, R. and Zhou, Z.-H. (2012), "Nyström method vs random fourier features: A theoretical and empirical comparison", In: *Advances in Neural Information Processing Systems*, pp. 476-484.

CC