Available online at www.sciencedirect.com
ScienceDirect
Mechanical Systems and Signal Processing 22 (2008) 155–171
Mechanical Systems and Signal Processing
www.elsevier.com/locate/jnlabr/ymssp

Multivariate statistics process control for dimensionality reduction in structural assessment

L.E. Mujica[a,*], J. Vehí[a], M. Ruiz[a], M. Verleysen[b], W. Staszewski[c], K. Worden[c]

[a]Department of Electronics, Computer Science and Automatic Control, University of Girona (UdG), 17071 EPS-PIV Campus Montilivi, Girona, Spain
[b]Electricity Department, Université Catholique de Louvain, Place du Levant, B-1348 Louvain-la-Neuve, Belgium
[c]Department of Mechanical Engineering, University of Sheffield, Mappin Street S1 3JD Sheffield, UK

Received 27 July 2006; received in revised form 2 May 2007; accepted 4 May 2007
Available online 17 May 2007

Abstract

This paper presents advantages of using techniques like principal component analysis (PCA), partial least square (PLS) and some extensions called multiway PCA (MPCA) and multiway PLS (MPLS) for reducing dimensionality in damage identification problem, in particular, detecting and locating impacts in a part of a commercial aircraft wing flap. It is shown that applying MPCA and MPLS is convenient in systems which many sensors are monitoring the structures, because the reciprocal relation between signals is considered. The methodology used for detecting and locating the impact uses the philosophy of case-based reasoning, where single PCA and PLS are used also for organizing previous knowledge in memory. Sixteen approaches combining those techniques have been performed. Results from all of them are presented, compared and discussed.

Keywords: Principal component analysis; Multiway principal component analysis; Partial least square; Multiway partial least square; Structural health monitoring

1. Introduction

Nowadays, laboratory instruments produce great quantities of data. This creates a data overload and usually a big amount of these data are wasted. The problem is to compress and/or to extract relevant information. Generally, there is a great deal of correlated or redundant information in procedure measures. This information must be compressed in a manner that retains the essential information and is more easily displayed than each of the variables individually. Also, essential information often lies not in any individual variable but in how the variables change with respect to one another, i.e. how they co-vary.

Dimensionality reduction is a way to transform vectors $X \varepsilon \mathbb{R}^d$ into new vector $T \varepsilon \mathbb{R}^q$, where $q < d$. This way must obey certain rules in order to be useful. Multivariate statistical process control (MPSC) can be considered as a tool to reduce the dimension of a data set [1]. MPSC techniques are a comparatively recent

*Corresponding author. Tel.: +34 972 418 487; fax: +34 972 418 259.
E-mail address: lemujica@eia.udg.es (L.E. Mujica).

0888-3270/$ - see front matter
doi:10.1016/j.ymssp.2007.05.001

development in engineering. The strength of these techniques lies in the ability to transform highly correlated, redundant and noisy data to a model whose components often offer insight into the underlying physical phenomena and relationships involved in the system. MPSC is based on two techniques: principal component analysis (PCA) [2] and partial least squares (PLS), also known as projection to latent structures [3].

In structural health monitoring (SHM), the data extracted are often unsuitable due to redundancy, correlation or large feature space. Consequently, dimensionality reduction methodologies have been the why and wherefore of some works [4]. Recently, techniques based on multivariate statistics [5,6] and statistical process control (SPC) [7] have been applied in structural damage detection. On the other hand, PCA has been used for performing data compression prior to the feature extraction process when data from multiple measurement points are available in order to enhance the discrimination between features from the undamaged and damaged structures [8]. This process transforms the time series from multiple measurement points into a single time series. Visualization and dimension reduction were implemented using PCA for damage detection [9]. PCA technique was used for condensing the frequency response functions data and their projection onto the most significant principal components, which were used as the artificial neural network input variables [10]. Moreover, PCA has also been used recently for several purposes including model reduction [11], dynamic characterization [12], sensor validation [13,14], modal analysis [15], parameter identification [16] or damage detection [17,18]. Some nonlinear extensions of PCA has been also used for SHM purposes [19,20].

It is not uncommon for SHM to have tens of sensors measuring variables for a long time and any change in the response of a single sensor is reproduced as well in the whole sensor response set. In order to study the relationship among all variables at any one time and its history, extensions of PCA and PLS are used; these extensions are known as multiway principal components (MPCA) and multiway partial least square (MPLS) [21].

In the community of SHM, PCA has been used mainly to extract features or reduce the dimensionality. This work proposes to use not only PCA to do that, but also to use PLS and its extensions (MPCA and MPLS) studying the correlation between sensors. This analysis is performed applying these techniques in combination with wavelet transform and case-based reasoning (CBR) [22] in order to detect and localizate impacts on an aircraft composite structure.

In SHM is well known that try to identify what are happening in the structure directly from the time series signal is not an easy task, for these reasons, researcher have studied many techniques which pretend extract more information from these signals. Among these, transformation methods (as Fourier, Wavelet, Hilbert, etc.) are considered good way for changing the point of view of the problem and see beyond [4,23]. In addition, combinations of these transformations with other techniques provide better results. Here, the combination of wavelet transform with MSPC is analysed to determine its usefulness.

On the other hand, CBR methodology demands the integration of another technique to classify and organize in memory old cases in such a way that the retrieve step could be fast and reliable. From another point of view, CBR has the need of reducing the dimension even more than the dimension obtained in the previous step. Therefore, this work propose also to apply single PCA and PLS so as to get this purpose.

The objective of the paper is to illustrate several dimensionality reduction techniques for SHM applications. Impact damage detection problem in an aircraft structure is used as an example. The structure, sensors, impact method and the experimental setup are explained with details in Section 2. Basic definitions about MSPC are presented in Section 3. Dimensionality reduction methodologies for locating impact identification are shown in Section 4, which describes the damage identification approach, techniques to reduce dimensionality and organization of the existing knowledge. Finally, results and conclusions are presented emphasizing the advantages and drawbacks of using MSPC in SHM.

2. Experimental procedure: an aircraft structure

Invisible and barely visible impact damage is a serious problem in the aerospace industry which affects the safety and service life of military and commercial aircrafts; it is considered to be the most common and important damage form in aircraft structures [24]. The use of composite materials in engineering structures is not only a great opportunity but also a major challenge to inspection and maintenance. Impacts with ground support equipment are one of the major causes of in-service damage to composite aerospace structures. Thus impact detection on composite structures has direct relevance to the problem of damage detection in these structures. Previous work in this area includes the following examples [25–29] among others.

This section describes the experimental setup which has been previously used showing that the problem of identifying impact locations can be solved successfully using neural networks, regression techniques and CBR locating impact positions to within a fairly accurate distance [30,31]. Figs. 1–4 are reproduced from [31]. Furthermore, the impact method and the data collecting way are presented in this section.

2.1. Wing flap structure

The structure used in this work is a section of a commercial aircraft wing flap with approximate dimensions shown in Fig. 1. This structure can be regarded as a small-scale version of part of a wing span with the corresponding features being a leading edge and trailing edge.

This structure is obviously more complex than a flat panel with a degree of curvature stretching over the surface of the structure. To impart strength into important areas of the panel there exist various ribs, spars and stringers running throughout the structure as well as the use of honeycomb cores at the leading and trailing edges. The trailing edge is composed of aluminium skins with an aluminium honeycomb core, the leading edge of composite skins with a light weight honeycomb core and the central section of thin composite material. Unfortunately, due to the nature of the origin of the wing flap section (being from a commercial aircraft) little is known about the specific materials and design parameters constituting the structure, such as the lay up of the composite.

The thickness of the composite material, as measured, used for the central section and the skins of the leading edge is 1.5 mm and that of the aluminium skin on the trailing edge 1 mm. The entire length of the panel taking into account the curvature of the surface is 1015 mm and its width 720 mm. As well as the ribs and stringers illustrated in Fig. 1 there is also a rib running along the middle of the central section of the panel. The locations of the major ribs are indicated by the positioning of rows of rivets along their lengths placed at fairly equally spaced intervals of 30 mm.

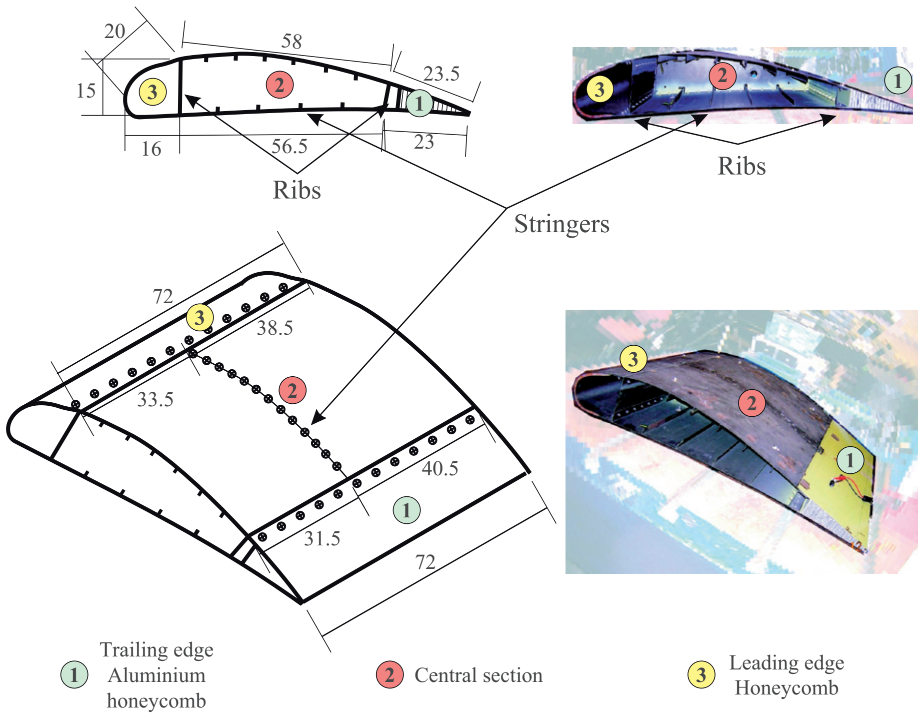

Fig. 1. Schematic representation of the wing flap section. Dimensions in cm.

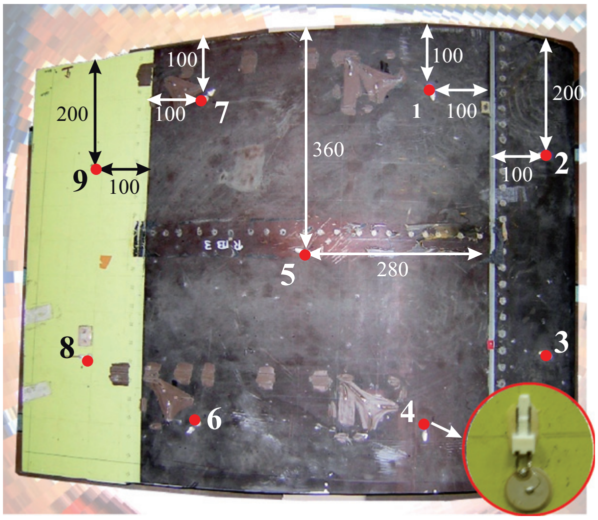

Fig. 2. Sensor locations (approximate measurements in mm).

Fig. 3. Wing panel with superimposed impact grid.

Nine low-profile, surface-bounded piezoceramic sensors (PIC 155, 10 mm in diameter and 1 mm thick) are used to measure impact strain data. The sensors and connectors are distributed over the surface of the flap; two on the leading edge, two on the trailing edge and five in the central section as shown in Fig. 2. They are attached to each other using short (to reduce effects of unwanted noise superimposed on the signals) wires soldered to the positive and negative elements of the sensors. The health of each sensor is monitored before testing to check that all the connections are substantial and signals are being received.

2.2. Impact method

Using datum lines drawn in the x and y directions a grid of approximately 60 mm $\times$60 mm is drawn on the trailing edge and the central section of the panel. Due to the importance of the leading edge with regard to a higher probability of impacts during flight from sources such as bird strikes, the grid is reduced in size on this section to 40 mm $\times$60 mm. The surface test area shown in Fig. 3 is impacted using a rubber tipped PCB instrumented hammer that is chalked before impact to mark the exact position of the impacts.

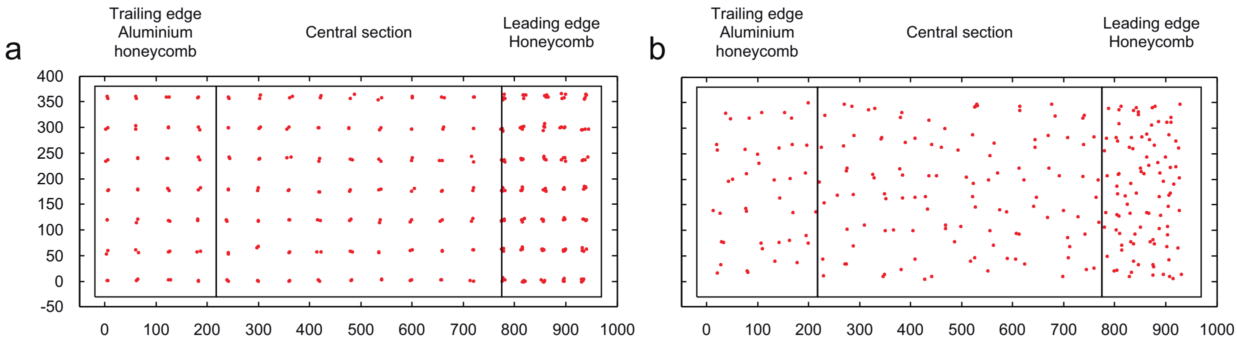

Fig. 4. Impact locations measured along the curvature of the structure. (a) Grid impacts, (b) random impacts.

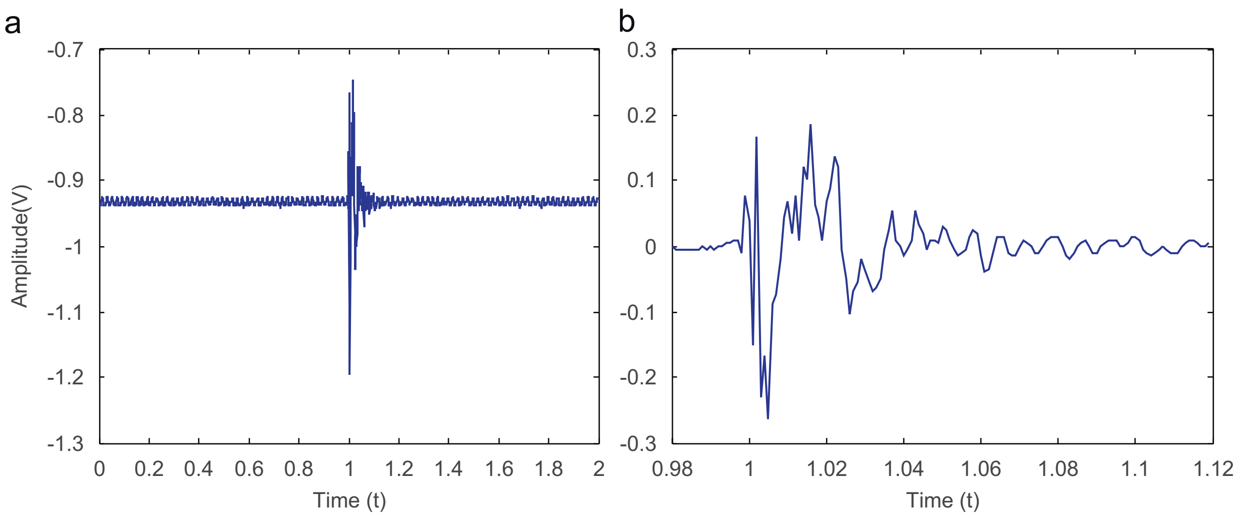

Fig. 5. Signal pre-processing example. (a) Sensed signal, (b) final signal.

Two different sets of data are taken: on the grid intersections and in a random array way over the entire test area. More impacts are repeated on the leading edge, again due to the greater interest in this area. The response of all nine sensors on the structure is recorded for a total of 574 impacts, the positions of which are measured from the x and y datum lines. Fig. 4 shows the distribution of the measured impacts over the test area.

2.3. Data collecting

The signals obtained during testing contain a lot of undesirable information, including: noise, different levels of unavoidable offset, trends, large amount of data-points, etc. (see Fig. 5a). The offsets are removed using the mean value of each signal. Furthermore, signals are cut off, eliminating data-points which does not contain any information (see Fig. 5b). Finally, the set of these signals is arranged in a three-dimension (3D) in which $j = 1, 2, \ldots, J$ sensors are recorded at $k = 1, 2, \ldots, K$ time instants throughout a particular experiment. Similar data are generated for a number of such experiment runs $i = 1, 2, \ldots, I$. That generates a *three-way data array* $\underline{X}\,\varepsilon\mathbb{R}^{IxJxK}$ as is illustrated Fig. 6a, where the height gives the number of experiments I, the width gives the number of time instants K, and the length gives the number of measurements (sensors) J. In this way, each frontal slice is a two-dimensional (2D)-matrix X which represents all measurements in one sensor. On the other hand, the position of the impact (x-location and y-location) is stored in a matrix Y as can be seen from Fig. 6b.

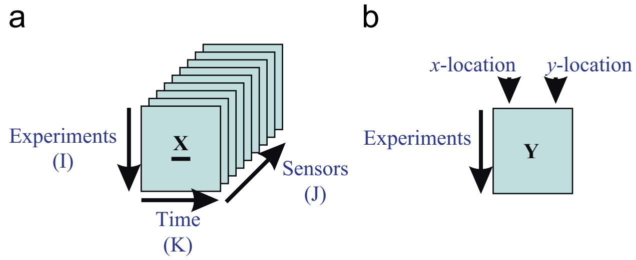

Fig. 6. Collected data: (a) Signals (b) impact location.

3. Linear dimension reduction: multivariate statistical process control

MSPC is a set of mathematical tools that can be used to extract information from a large amount of data [32]. The basis of MSPC schemes are the statistical projection techniques of PCA [33] and PLS [3]. In this section, a brief description of these techniques and its extensions known as MPCA and MPLS [34] are presented.

3.1. Principal component analysis

PCA is a standard tool for data compression and information extraction which finds combinations of variables or factors that describe major trends in a data set [35]. PCA is concerned with explaining the variance–covariance structure through a few linear combinations of the original variables. Its general objectives are data reduction and interpretation. The multivariate data are organized in K variables and I samples per variable. In this work, PCA is applied to each sensor separately (2D-matrix $\underline{X}$); on the other hand, *time* is treated as independent variable which its dimension will be reduced.

The first step in applying PCA is to standardize the data matrix X, since PCA is scale variant [36]. Several studies of scaling are presented in literature: continuous scaling (CS), group scaling (GS) and autoscaling (AS) [37,38]. According to these studies, GS is selected for this work because it considers changes between variables and does not process independently the variables. The mean trajectories are removed and all variables are made to have equal variance. As a consequence, the experiment trajectories of the sensors and their standard deviations, often non-linear in nature, are removed from the data.

Once the variables have been standardized, the covariance matrix S is calculated:

$$S = \frac{1}{I-1} X^{\mathrm{T}} X. \tag{1}$$

The matrix $\hat{P}$ where columns are the eigenvectors of S and the diagonal matrix λ with eigenvalues of S on the main diagonal are found:

$$S\hat{P} = \hat{P}\lambda. \tag{2}$$

Each eigenvalue is associated to an eigenvector. The eigenvector with the highest eigenvalue represents the most important pattern in the data, i.e. contains the largest quantity of information, therefore this vector is called the *principal component* of the data set. Ordering the eigenvectors by eigenvalue, highest to lowest, gives the components in order of significance. In order to reduce the dimensionality, the less important components can be eliminated (information is lost, but if the eigenvalues are small, this information is not much), then only the n first eigenvectors are chosen (loading vectors and denoted by P) and the final data set will be n-dimensional. The projected matrix T (or score vectors) in the new space is defined by

$$T = XP, \tag{3}$$

and the projection of T back onto the K-dimensional observation space is

$$\hat{X} = TP^{\mathrm{T}}. \tag{4}$$

The difference between X and $\hat{X}$ is the residual matrix E [32]:

$$X = \hat{X} + E, \tag{5}$$

$$X = TP^{\mathrm{T}} + E. \tag{6}$$

3.2. Partial least square

PLS, also known as *projection to latent structures* [3], computes loading P and score vectors T that are correlated with the predicted block Y while describing a large amount of the variation in the predictor matrix X. Here, $X \varepsilon \mathbb{R}^{I \times K}$ and it is defined as in PCA according to Fig. 6a and $Y \varepsilon \mathbb{R}^{I \times H}$ where H is the number of observations and I the number of samples. In this work, Y is defined by the impact location, it means (as can be seen from Fig. 6b) that $H = 2$ and I is the number of experiments. The PLS is achieved by decomposing X and Y into a combination of loadings P and Q (formed by orthogonal vectors), scores T and U (the projections of the loading vectors associated with the first singular values) and residual matrices E and F [39]:

$$X = TP^{\mathrm{T}} + E, \tag{7}$$

$$Y = UQ^{\mathrm{T}} + F. \tag{8}$$

Firstly, the input and output variables are projected onto a subspace of orthogonal principal components, giving the input and output scores T and U, before an ordinary least square regression is carried out between each pair of corresponding input and output scores. The heart of PLS methodology is the nonlinear iterative partial least square (NIPALS) algorithm [3].

3.3. Multiway principal component analysis

Although in SHM is not unusual to measure a lot of variables, this does not always mean that the same quantity of independent things are taking place. The measured variables are autocorrelated in time and extremely highly correlated with one another at any give time. Moreover, the relationship among all the variables at any one time is not only important, but also the whole past history of these variables. MPCA is used to compress such data to extract the information by projecting the data into a low-dimensional space condensing both the variables and their time history [34]. MPCA is equivalent to performing ordinary PCA on a large two-dimensional (2D) matrix constructed by unfolding the 3D matrix. Six possible ways of unfolding have been suggested before [37]: they are indicated in Table 1, showing the structure of the unfolding matrix and the direction that remains unaltered.

When aiming at PCA-based monitoring, B- and D- unfolding will lead to models that are equivalent to the models constructed on the C- and E- unfolded matrices, respectively. In this article, the intention is to summarize both sensor and time information, and to keep the experiments unaltered; therefore E-unfolding will be applied. In Fig. 7 it can be seen how experiment unfolding is made. Each frontal slice represents all measurements of one sensor. Slices are put next to each other; one row represents the data of one experiment. Afterward, PCA is applied to this unfolded matrix.

Table 1
Ways of unfolding a three-way data matrix, according to [37]

Type	Structure	Direction
A	$IK \times J$	Sensor
B	$JI \times K$	Time
C	$IJ \times K$	Time
D	$I \times KJ$	Experiment
E	$I \times JK$	Experiment
F	$J \times IK$	Sensor

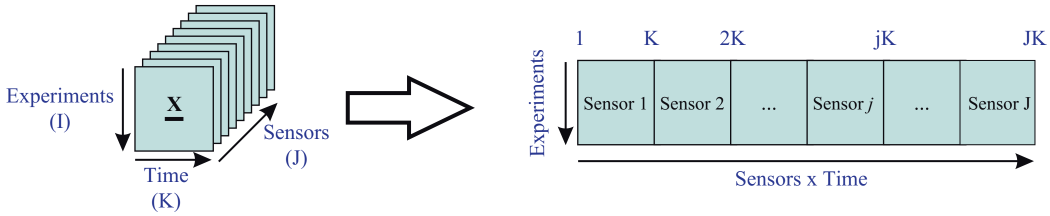

Fig. 7. Decomposition of X to 2D ($I \times JK$).

The previous process allows decomposing the three-way array $\underline{X}$ into a series of principal components consisting of score vectors T and loading matrices P, plus a residual 3D-matrix $\underline{E}$, in this way and in accordance with the principles of PCA, it separates the data in two parts:

$$\underline{X} = T \bigotimes P + \underline{E}, \tag{9}$$

where the first part or systematic part consists in a Kronecker product or Tensor product between T (related to the experiment) and P (related to the sensors and their time variation) [34].

3.4. Multiway partial least square

The same multivariate SPC monitoring ideas that were developed using MPCA can be extended directly using MPLS when predicted data Y are available. Comparing MPCA and MPLS, MPCA only makes use of the sensor measurements $\underline{X}$, taken throughout the duration of the experiment and is focused in its variance, while MPLS focuses on the variance of $\underline{X}$ and Y.

MPLS is an extension of PLS to handle data in 3D arrays. MPLS is equivalent of performing ordinary PLS on a large 2D-matrix X formed by unfolding the three-way array $\underline{X}$ in one of the six possibilities cited in Table 1 [40]. Due to the reasons previously explained, E-unfolding is chosen.

4. Dimensionality reduction for locating impact damage in structures

On the whole, the methodology applied in this work for locating impacts using CBR [41,42] consists of two principal tasks: (i) Case defining or feature extraction which its goal is calculate indicators from the data that can be used to represent the impacts and (ii) building the casebase, which means to recognize patterns of these cases and classify them in order to improve the matching into the CBR system. In both tasks, the problem can be considered as a dimensionality reduction problem. In this section, the employment of dimensionality reduction techniques for locating impacts are studied and evaluated.

4.1. Impact location identification using case-based reasoning

In a "learning mode" the first group of data (grid impacts) is used to generate a set of cases. In principle, each case is defined by the impact location and its dynamic response. Due to the large amount of data, the dynamic response information should be compressed and extracted, therefore, its dimension is reduced. The casebase is an array in memory organizing all the cases to facilitate the search for the cases most similar to the current problem. These casebases are built using the first set of cases (322 grid impacts) which can be seen in Fig. 4 and they will be used in diagnosing future situations by analogy (see Fig. 8).

When the system is in the "operation mode" the data group of random impacts shown in Fig. 4b are used to test and verify the accuracy of the methodologies. In each new case, several cases are retrieved from the casebase. In the adapting step, it is decided how many cases by sensors are used to estimate the location of the impact. Finally, each new experience is retained once the damage has been detected. In order to evaluate its accuracy, the radial error described by Eq. (10) and Fig. 9 is calculated by experiment

$$radial_{\mathrm{error}} = \sqrt{(x_{\mathrm{real}} - x_{\mathrm{detected}})^2 + (y_{\mathrm{real}} - y_{\mathrm{detected}})^2}. \tag{10}$$

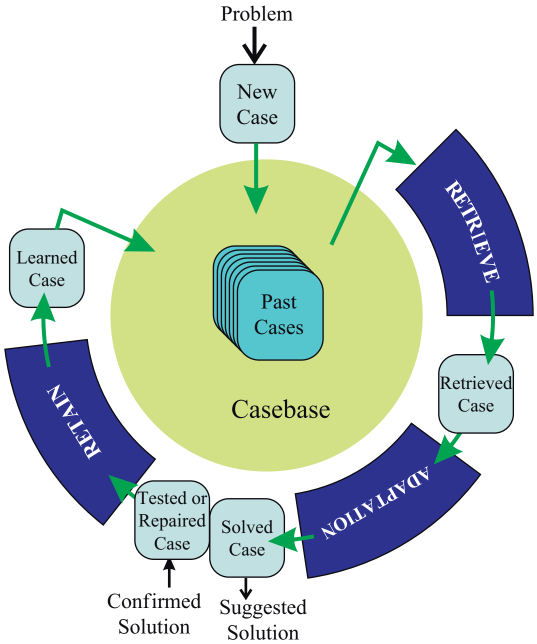

Fig. 8. Case-based reasoning cycle.

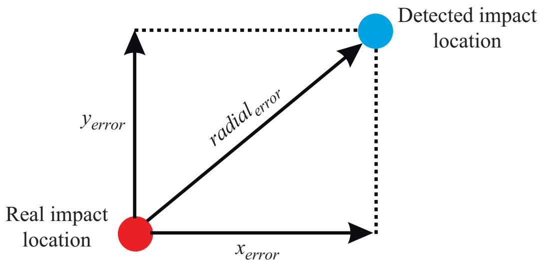

Fig. 9. Graphical representations of error definitions.

4.2. Cases defining

The cases defined by the minimal representation of the dynamic response to the impact are obtained applying several combinations of MSPC techniques in order to reduce the dimensionality of the collected data shown in Figs. 6 and 10a. The analysis is performed following the next steps (as can be seen from Fig. 10b):

- Either using directly the time signals or its wavelet coefficients.
- Either analysing each sensor separately (dividing the 3D matrix in several 2D matrices) or considering the whole sensor response set (unfolding the 3D matrix data in one 2D matrix).
- Either applying PCA or PLS (using the impact position as predicted matrix Y).

In this way, eight methodologies are analysed:

1. *Wavelet* and *PCA*: PCA applied to wavelet coefficients.
2. *Wavelet* and *PLS*: PLS applied to wavelet coefficients.
3. *PCA*: PCA applied to time signal.
4. *PLS*: PLS applied to time signal.
5. *Wavelet* and *MPCA*: MPCA applied to wavelet coefficients.

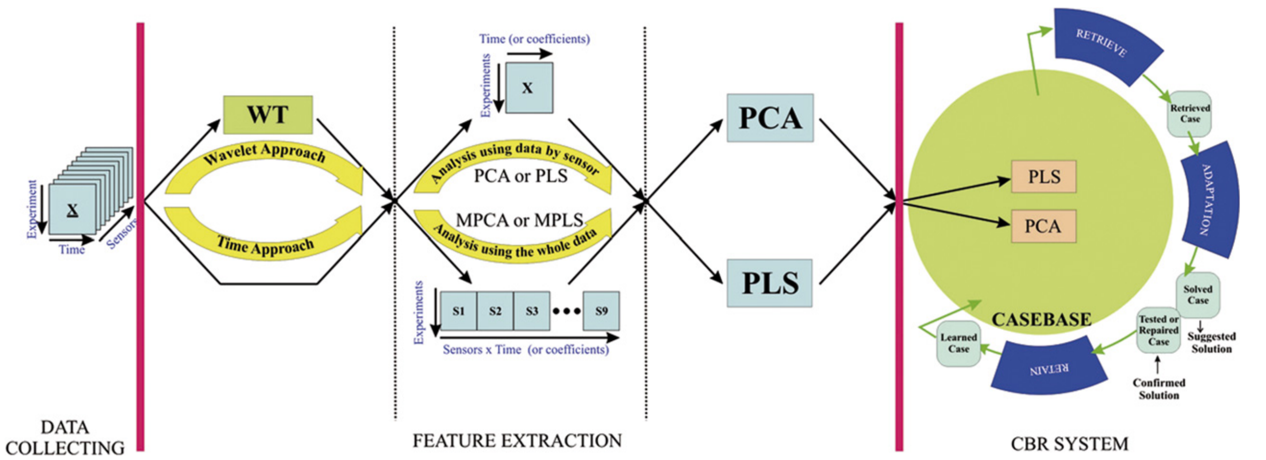

Fig. 10. Summary of dimensionality reduction methodologies.

6. *Wavelet* and *MPLS*: MPLS applied to wavelet coefficients.
7. *MPCA*: MPCA applied to time signal.
8. *MPLS*: MPLS applied to time signal.

The first four methodologies, PCA and PLS are applied to each sensor separately and one casebase is built for each configuration. On the last four, MPCA and MPLS are applied to the 3D matrix data in order to consider reciprocal relations between sensors and to build only one casebase for all sensors.

Finally, the set of cases required to build the casebase in the impact location identification using CBR is defined by the projected matrix T (or score vectors that in this instance are the minimal representation of the impact response) and the location of this impact (Y).

4.3. Casebase building

The set of cases, which have been achieved using all previously described methodology, is classified and organized in memory for recovering at the required time. Among other possibilities for organization, projecting the cases to a *low*-dimension is an effective approach. In this way, the retrieve of the stored cases set with similar characteristics is an easy task by mean of classical Euclidean distance. In order to compare and evaluate the performance of the proposed linear methodologies to reduce dimensionality in impact location identification, two kind of casebases have been built combining also PCA and PLS (see Fig. 10c). Thus, its efficiency like pattern classifier tool is evaluated as well.

4.3.1. Using principal component analysis

Projecting again by means of PCA the set of cases to a new space 2D, every case can be represented by a point as is shown in Fig. 11. In this way, a new impact is projected as well to this space and the level of similarity with every stored case can be calculated. As can be seen from the figure, the mark of the every case is different according to the position of the impact and it is fairly clear that a cluster between wing sections exists regardless of these groups overlap. The same concept can be applied using not only projecting to 2D but also to 3D or even 10D, but of course, with the disadvantage that it is difficult or impossible to show by mean of a graphic.

4.3.2. Using partial least square

Following the procedure previously described but applying PLS to project to a new space 2D and using the impact locations as predicted matrix Y, a casebase can be organized as well. Fig. 12 shows the stored cases and its arrangement. Projections 3D or 10D are possible too.

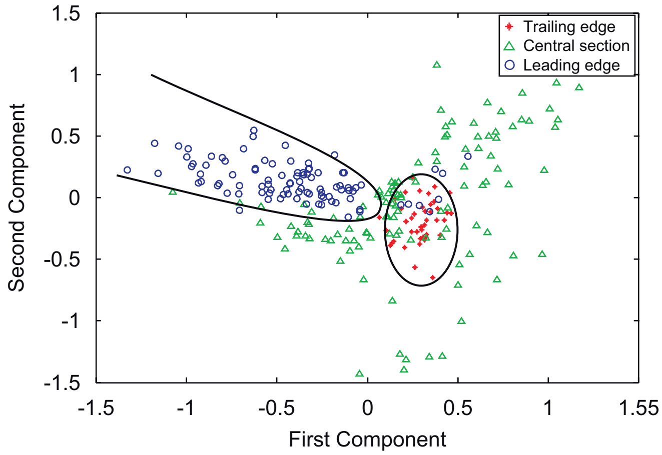

Fig. 11. Casebase representation using PCA.

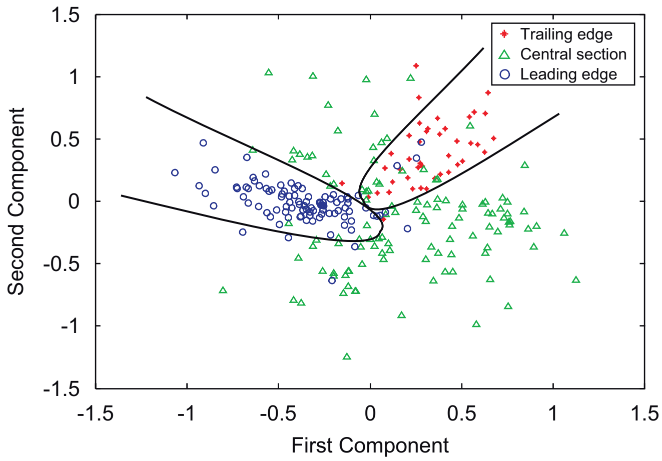

Fig. 12. Casebase representation using PLS.

5. Analysis results and discussion

Since the objective of this work is to illustrate the application of the MSPC techniques in SHM, specifically for reducing the dimensionality and the hybridization with CBR for organizing the previous knowledge in memory, a combination of these techniques has been performed in both tasks. Therefore, in some cases, PCA and partial PLS are applied twice. MPCA and MPLS are used for reducing the dimensionality, but not in the second task.

A total of 16 approaches have been performed combining: (i) Eight methodologies for reducing dimensionality: from 140 samples by signal to 10 and 20 datapoints by signal (described in Section 4.2) and (ii) simple PCA and PLS for organizing casebases using 2, 3 and 10 principal components (detailed in Section 4.3). In Table 2, the average radial error (see Fig. 9 and Eq. (10)) is shown for every approach.

Analysing firstly the casebases, it is clear that 2 or 3 principal components using PCA or PLS do not provide an enough accuracy to identify the location of the impact, regardless of the methodology used for reducing dimension.

Choosing casebase using PCA and PLS with 10 components, a graph shown in Fig. 13 is created. From this figure, the average radial error in the selected approaches can be compared. At first sight, it is clear that using the set of approaches which apply MSPC, the error is a little larger than with the other approaches.

Table 2
Average radial error (mm) using all signals

No.	Method	PC	Casebase					
			PCA			PLS		
			2PC	3PC	10PC	2PC	3PC	10PC
1	Wavelet and PCA	10	97,71	69,89	52,61	96,39	66,32	50,87
		20	97,71	69,89	52,61	86,06	66,80	49,44
2	Wavelet and PLS	10	93,67	69,20	50,75	87,33	68,96	51,63
		20	93,64	69,20	49,66	86,49	68,98	51,49
3	PCA	10	97,98	70,58	52,56	97,84	66,17	50,41
		20	97,98	70,58	52,56	85,69	67,63	49,43
4	PLS	10	93,16	69,08	50,68	86,18	68,53	51,68
		20	93,16	69,08	49,56	85,38	68,27	50,86
5	Wavelet and MPCA	10	172,88	106,54	59,92	186,83	142,08	58,64
		20	172,88	106,54	59,92	153,35	109,73	64,01
6	Wavelet and MPLS	10	146,49	104,66	61,03	126,39	96,48	63,04
		20	146,49	104,66	61,03	116,55	95,02	64,86
7	MPCA	10	172,88	106,54	59,92	186,56	142,08	58,64
		20	172,88	106,54	59,92	153,76	109,75	64,01
8	MPLS	10	155,22	106,93	61,28	133,63	101,57	62,88
		20	155,22	106,93	61,28	124,36	99,96	65,34

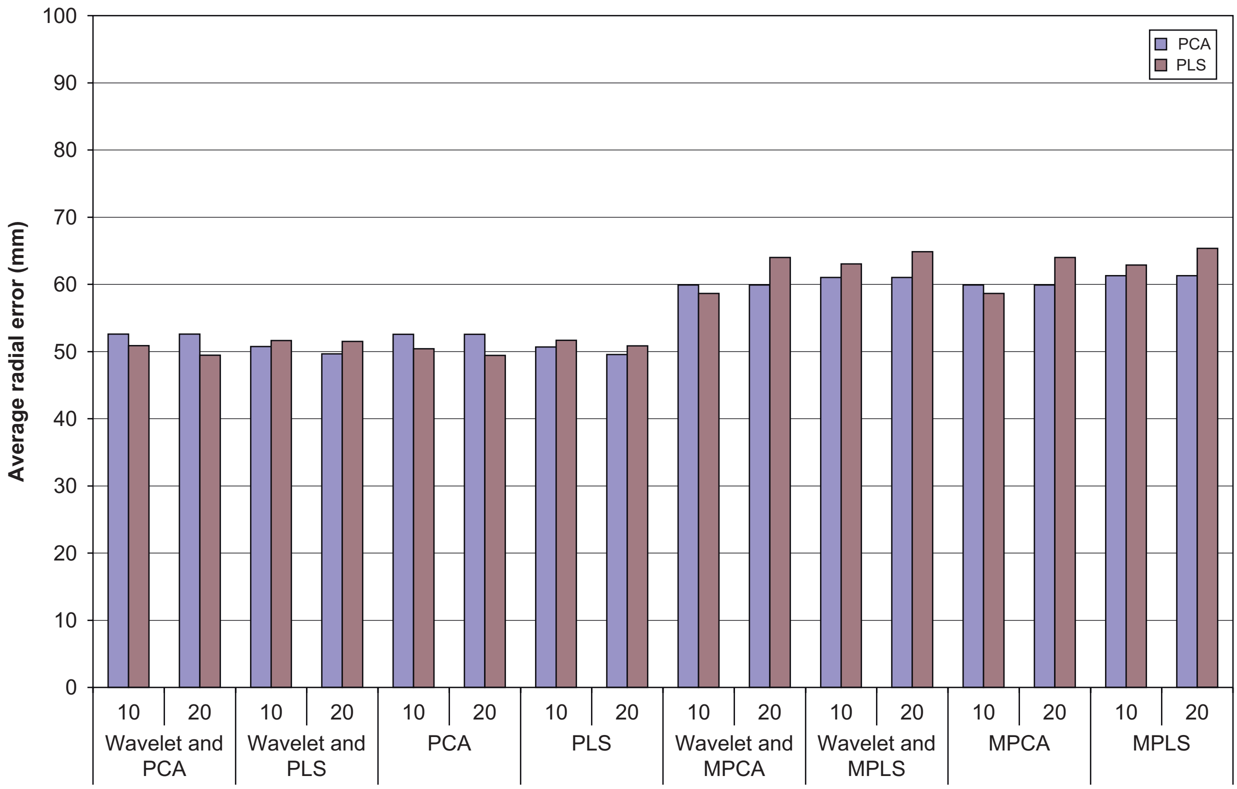

Fig. 13. Average error by methodology using the whole surface.

Table 3
Average radial error (mm) separating signals by wing sector

No.	Method	PC	Trailing edge		Central section		Leading edge	
			PCA	PLS	PCA	PLS	PCA	PLS
1	Wavelet and PCA	10	55,36	51,91	61,78	65,77	39,68	36,20
		20	55,36	54,03	61,78	62,33	39,68	34,98
2	Wavelet and PLS	10	52,09	52,19	60,98	59,93	35,79	31,66
		20	51,43	51,99	60,98	53,94	35,55	32,28
3	PCA	10	55,36	51,91	61,53	65,29	39,65	36,16
		20	55,36	53,79	61,53	62,61	39,65	35,16
4	PLS	10	51,85	52,30	60,50	60,22	35,94	31,06
		20	51,32	51,80	61,52	55,28	35,67	32,21
5	Wavelet and MPCA	10	60,41	60,74	72,28	77,24	40,97	39,22
		20	60,41	56,13	72,28	68,82	40,97	38,13
6	Wavelet and MPLS	10	53,85	50,80	63,40	64,90	37,90	35,15
		20	52,77	47,52	63,11	62,25	37,57	35,26
7	MPCA	10	60,81	60,74	71,21	78,48	41,02	39,04
		20	60,81	56,32	71,21	68,53	41,02	37,62
8	MPLS	10	53,82	50,36	62,59	65,54	37,44	35,37
		20	52,68	48,89	62,24	61,63	37,41	34,18

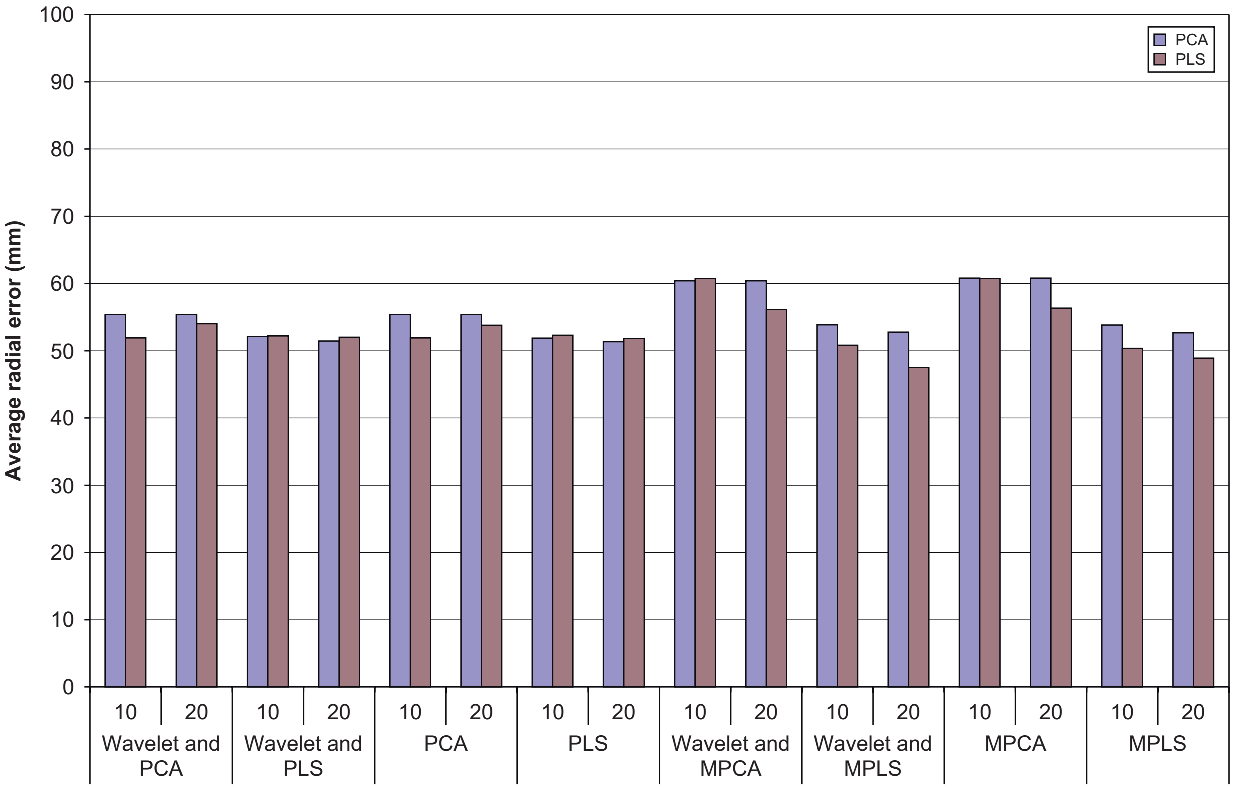

Fig. 14. Average error by methodology using the trailing edge section.

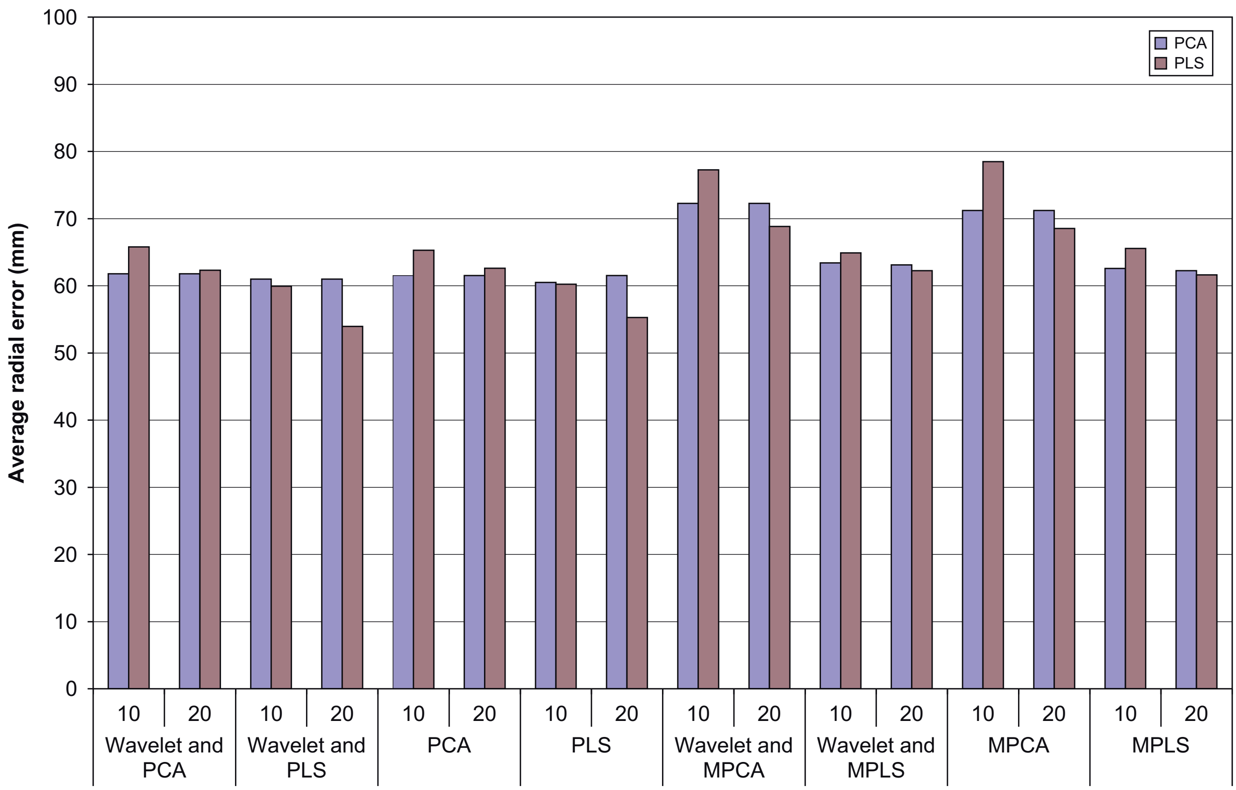

Fig. 15. Average error by methodology using the central section.

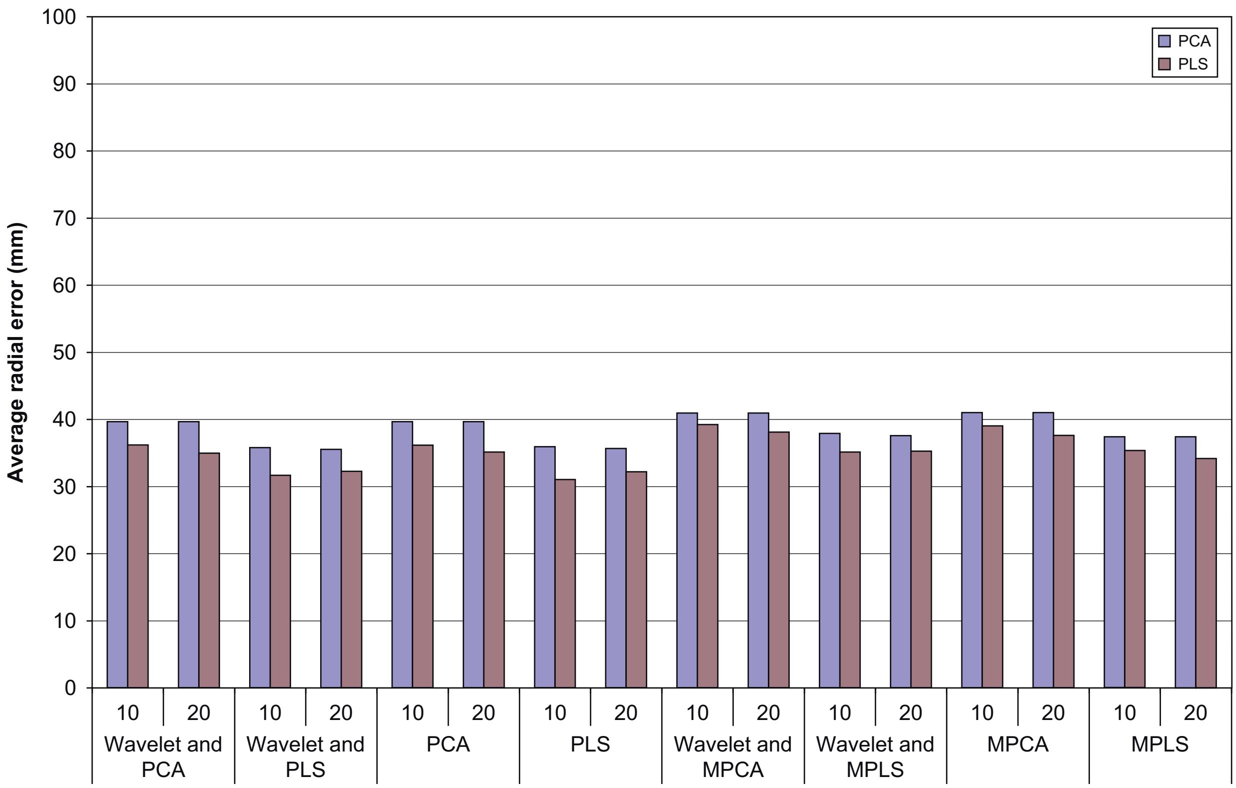

Fig. 16. Average error by methodology using the leading edge section.

However, considering that with these methodologies only one casebase is trained (and not nine as in the first four methodologies), the computational cost is very small. If the system has hundreds of sensors, there is a clear advantage.

The experimental setup (wing flap) is a very complex structure and has three sections perfectly distinguishable; they are entirely different in shape and materials, and the reciprocal relation between dynamic responses caused by the impacts cannot exist or can be very small. The main idea of using MPCA and MPLS is focused to consider that relation. For that reason, authors consider that the benefits of using any of those methodologies (illustrated in Fig. 10) is not perceived. Therefore, the wing sections are analysed separately. In other words, if the trailing edge is studied, impacts produced there and signals collected by sensors 8 and 9 (located in this area) are analysed as a system, independently of the rest of the structure. In this way, the relation between dynamic responses can be observed and studied.

As can be seen from the Table 3 and Figs. 14–16, the average radial error is nearly the same in every approach. Given this, it may be inferred that advantages of using MPCA and MPLS in identification impact location have been shown out. Those advantages can be expanded to SHM in general.

Another point important to highlight is to compare the average radial error in all the three sections. The accuracy in the leading edge section is much better than the other ones. That is because more experiments have been performed there due to the importance of the this section with regard to a higher probability of impacts during flight from sources such as bird strikes. Therefore, the more experiments are carried out, the more robustness.

6. Conclusions

This paper shows the advantages of using multivariate statistical process control (MSPC) for dimensionality reduction in structural health monitoring (SHM); the example of detecting impacts in a part of wing aircraft is studied. The surface of the wing is impacted and its response is measured by nine sensors. Grid impacts were used to train the casebase into a CBR methodology and random impacts for testing it. MSPC techniques are used to reduce the dimensionality of the problem. MSPC techniques include: principal component analysis (PCA), partial least square (PLS), and extensions called multiway PCA (MPCA) and multiway PLS (MPLS). On the other hand, PCA and PLS are also applied as tool to recognize patterns and classify cases.

A total of 16 approaches have been tested using several combinations of four techniques used to reduce dimensionality including or not wavelet transform and two for organizing the casebases. Results show that the system is accurate, and demonstrate the feasibility of using MSPC in SHM. Moreover, multiway extensions of MSPC are very useful in systems that involves several sensors since it decreases drastically the computation cost: a single casebase is built for the whole system instead of one by sensor. From the results can be seen also that the contribution of the wavelet transform application is very few, almost nothing, in this way, it is shown that MSPC is useful by applied directly to signals, and could be not necessary to use another transformations (wavelet, FFT, Hilbert, etc.).

Acknowledgements

This work has been partially funded by the European Union (European Regional Development Fund) and the Spanish government through the coordinated research projects DPI2003-07146-C02-02 and DPI2004-07167-C02-02, by the government of Catalonia through 2005SGR-00296 and the University of Girona. The authors would like to thank the support to Gurkan Sin who has contributed with new ideas. We are grateful also to Ms Jenny LeClerc who has collected the data within her Master Thesis. Michel Verleysen is a Research Director of the Belgian National Fund for Scientific Research.

References

[1] R.A. Johnson, D.W. Wichern, Applied Multivariate Statistical Analysis. 0-130-41146-9, third ed., Prentice-Hall, Upper Saddle River, NJ, USA, 1988.
[2] L. Smith, A tutorial on principal components analysis, Technical Report, University of Otago, New Zealand, February 2002.

[3] P. Geladi, B. Kowalski, Partial least-squares regression: a tutorial, Analytica Chimica Acta 185 (1986) 1–17.
[4] W.J. Staszewski, C. Boller, G.R. Tomlinson, Health Monitoring of Aerospace Structures: Smart Sensor Technologies and Signal Processing, Wiley, Chichester, 2004.
[5] S.W. Doebling, C.R. Farrar, M.B. Prime, A summary review of vibration-based damage identification methods, The Shock and Vibration Digest 30 (2) (1998) 91–105.
[6] K. Worden, G. Manson, Damage identification using multivariate statistics: kernel discriminant analysis, Inverse Problems in Engineering 8 (2000) 25–46.
[7] M.L. Fugate, H. Sohn, R.C. Farrar, Vibration-based damage detection using statistical process control, Mechanical Systems and Signal Processing 15 (4) (2001) 707–723.
[8] H. Sohn, J.A. Czarnecki, C.F. Farrar, Structural health monitoring using statistical process control, Journal of Structural Engineering 126 (1) (2000) 1356–1363.
[9] K. Worden, G. Manson, Visualization and dimension reduction of high-dimensional data for damage detection, imac 17, in: S. Ganeriwalaand, S. Patel, H. Hartung (Eds.), Proceedings of the 17th International Modal Analysis Conference, 1999, pp. 1576–1585.
[10] C. Zang, M. Imregun, Structural damage detection using artificial neural networks and measured frf data reduced via principal component projection, Journal of Sound and Vibration 242 (5) (2001) 813–827.
[11] M.F.A. Azeez, A.F. Vakakis, Proper orthogonal decomposition of a class of vibroimpact oscillations, Journal of Sound and Vibration 240 (1998) 859–889.
[12] R. Kappagantu, B.F. Feeny, Dynamical characterization of a frictionally excited beam, Nonlinear Dynamics 22 (2000) 317–333.
[13] M.I. Friswell, D.J. Inman, Sensor validation for smart structures, Journal of Intelligent Material Systems and Structures 10 (1999) 973–982.
[14] G. Kerschen, P. De Boe, J.C. Golinval, K. Worden, Sensor validation using principal component analysis, Smart Materials and Structures 14 (2005) 36–42.
[15] S. Han, B.F. Feeny, Application of proper orthogonal decomposition to structural vibration analysis, Mechanical Systems and Signal Processing 17 (2003) 989–1001.
[16] V. Lenaerts, G. Kerschen, J.C. Golinval, Proper orthogonal decomposition for model updating of non-linear mechanical systems, Mechanical Systems and Signal Processing 15 (2001) 31–43.
[17] P. De Boe, J.C. Golinval, Principal component analysis of piezo-sensor array for damage localization, Structural Health Monitoring: An International Journal 2 (2) (2003) 137–144.
[18] A.M. Yan, G. Kerschen, P. De Boe, J.C. Golinval, Structural damage diagnosis under varying environmental conditions. Part i: a linear analysis, Mechanical Systems and Signal Processing 19 (4) (2005) 847–864.
[19] H. Sohn, K. Worden, C.R. Farrar, Statistical damage classification under changing environmental and operational conditions, Journal of Intelligent Material Systems and Structures 13 (9) (2002) 561–574.
[20] A.M. Yan, G. Kerschen, P. De Boe, J.C. Golinval, Structural damage diagnosis under varying environmental conditions. Part ii: local pca for non-linear cases, Mechanical Systems and Signal Processing 19 (4) (2005) 865–880.
[21] S. Wold, P. Geladi, K. Esbensen, J. Öhman, Multiway principal component and pls analysis, Journal of Chemometrics 1 (1987) 41–56.
[22] R. Lopez de Mantaras, E. Plaza, Case-based reasoning: an overview, AI Communications 10 (1) (1997) 21–29.
[23] W.J. Staszewski, Wavelet based compression and feature selection for vibration, Journal of Sound and Vibration 211 (5) (April 1998) 735–760.
[24] W.J. Staszewski, Monitoring on-line integrated technologies for operational reliability-monitor, Air & Space Europe 2 (4) (July 2001) 67–72.
[25] J. Haywood, P.T. Coverley, W.J. Staszewski, K. Worden, An automatic impact monitor for a composite panel employing smart sensor technology, Smart Materials and Strctures 14 (2005) 265–271.
[26] W.J. Staszewski, K. Worden, R. Wardle, G.R. Tomlinson, Fail-safe sensor distributions for impact detection in composite materials, Smart Materials and Structures 9 (2000) 298–303.
[27] K. Worden, W.J. Staszewski, Impact location and quantification on a composite panel using neural networks and a genetic algorithm, Strain (36) (2000) 61–70.
[28] R.T. Jones, J.S. Sirkis, E.J Friebele, A.D. Kersey, Location and magnitude of impact detection in composite plates using neural networks, in: W.B. Spillman (Ed.), Proceedings of SPIE Vol. 2444, Smart Structures and Materials 1995: Smart Sensing, Processing, and Instrumentation, April 1995, pp. 469–480.
[29] P.T. Coverley, W.J. Staszewski, Impact damage location in composite structures using optimized sensor triangulation procedure, Smart Materials and Structures 12 (2003) 795–803.
[30] J. LeClerc, J. Haywood, W. Satszewski, K. Worden, Impact detection in an aircraft composite panel—a neural approach, in: C. Boller, W.J. Satszewski (Eds.), Proceeding of 2nd European Conference on Structural Health Monitoring, University of Sheffield, DEStech Publications, Inc., 2004, pp. 407–414.
[31] L.E. Mujica, J. Vehí, W. Staszewski, K. Worden, Impact damage detection in aircraft composites using knowledge-based reasoning, Proceeding of the Third European Structural Health Monitoring, 2006, pp. 601–608.
[32] E.L. Russell, L.H. Chiang, R.D. Braatz, Data-Driven Methods for Fault Detection and Diagnosis in Chemical Processes (Advances in Industrial Control), 1-85233-258-1, Springer, London, 2000.
[33] I.T. Jolliffe, Principal Component Analysis, Springer Series is Statistics, second ed., Springer, 2002.

[34] P. Nomikos, J.F. MacGregor, Monitoring batch processes using multiway principal component analysis, AIChE Journal 40 (8) (August 1994) 1361–1375.

[35] B.M. Wise, N.B. Gallagher, S. Watts, D.D. White, G.G. Barna, A comparison of pca, multiway pca, trilinear decomposition and parallel factor analysis for fault detection in a semiconductor etch process, Journal of Chemometrics 13 (1999) 379–396.

[36] E. Martin, J. Morris, S. Lane, Monitoring process manufacturing performance, IEEE Control Systems Magazine 22 (5) (October 1992) 26–39.

[37] J. Westerhuis, T. Kourti, J. MacGregor, Comparing alternative approaches for multivariate statistical analysis of batch process data, Journal of Chemometrics 13 (1999) 397–413.

[38] M. Ruiz, K. Villez, G. Sin, J. Colomer, P.A. Vanrolleghem, Influence of scaling and unfolding in PCA based monitoring of nutrient removing batch process, in: proceedings 6th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes (Safeprocess2006). Beijing, China, August 29–September 1, 2006 (on CD-ROM).

[39] A. Hoskuldsson, Pls regresion methods, Journal of Chemometrics 2 (3) (1988) 211–228.

[40] P. Nomikos, J.F. MacGregor, Multi-way partial least squares in monitoring batch processes, Chemometrics and Intelligent Laboratory Systems 30 (1995) 97–198.

[41] L.E. Mujica, J. Vehí, J. Rodellar, P. Kolakowski, A hybrid approach of knowledge-based reasoning for structural assessment, Smart Materials and Structures 14 (2005) 1554–1562.

[42] P. Kolakowski, L.E. Mujica, J. Vehí, Two approaches to structural damage detection: Vdm vs cbr, Journal of Intelligent Material Systems and Structures 16 (2006) 63–79.