Structural Health Monitoring Based on Principal Component Analysis: Damage Detection, Localization and Classification

D.A. Tibaduiza, L.E. Mujica, J. Rodellar

Abstract—Structural Health Monitoring is an area that its main objective is the verification of the state or the health of the structures in order to ensure proper performance and maintenance cost savings using non-destructive tests.

Currently, in the CoDAlab group it is used an active piezoelectric system which involves the use of piezoelectric transducers that are attached to the surface of the structure in order to apply vibrational excitations and collect dynamic responses at different points. As pattern recognition technique, Principal Component Analysis is used to perform the analysis, built a base-line model of the structure without damage and, subsequently to compare the data of the current structure under test. Different indices are calculated to determine how different is the structure under test. Using these indices, it is possible to detect, classify and locate the damage by means of the contribution of each sensor to each index. These methodologies are tested using two different structures, one aircraft turbine blade and one aluminium plate, which were instrumented with seven and four Piezoelectric transducer discs respectively. Seven simulated damages were made in the aircraft turbine blade and four real damages in the aluminium plate.

Index Terms—SHM, Principal Component Analysis (PCA), Damage Indices (DI), Damage Detection, Damage Localization, Damage Classification

1 INTRODUCTION

Structural Health Monitoring (SHM) is the integration of elements of actuation and sensing with different mathematics and computational techniques in order to know the health of a structure using non-destructive techniques. All the data obtained from the structure are analyzed to detect abnormal characteristics and to define the health of the structure. This is really important because this monitoring can define if the structure can work and in which conditions. Different benefits are derived from the implementation of SHM, some of them are: Knowledge about of the behavior of the structure under different loads and different environmental changes, Knowledge of the current state in order to verify the integrity of the structure and determine whether a structure can work properly or whether it need to maintain or replace, and, therefore maintenance cost saving. The paradigm of damage identification (comparison between the data collected from the structure without damages and the current structure in order to determine if there are any changes) can be tackled as a pattern recognition application.

2 PRINCIPAL COMPONENT ANALYSIS (PCA)

Principal Component Analysis is a technique of multivariable and megavariate analysis [5] which may provide arguments for how to reduce a complex data set to a lower dimension and reveal some hidden and simplified structure/patterns that often underlie it. The goal of PCA is to discern which dynamics are more important in the system, which are redundant and which are just noise [12]. This goal is essentially achieved by determining a new space (coordinates) to re-express the original data filtering that noise and redundancies based on the variance-covariance structure of the original data. PCA can be also considered as a simple, non-parametric method for data compression and information extraction, which finds combinations of variables or factors that describe major trends in a confusing data set [10]. Among their objectives it can be mentioned: to generate new variables that could express the information contained in the original set of data, to reduce the dimensionality of the problem that is studied, to eliminate some original variables if its information is not relevant. In order to develop a PCA model it is necessary to arrange the collected data in a matrix \( X \) this \( m \times n \) matrix contains information from \( n \) sensors and \( m \) experimental trials [7]. Since physical variables and sensors have different magnitudes and scales, each data-point is scaled using the mean of all measurements of the sensor at the same time and the standard deviation of all measurements of the sensor. Once the variables are normalized the covariance matrix \( C_x \) is calculated as follows:

\[
C_x = \frac{1}{m-1}X^T X
\]
\( \mathbf{C} \) is a square symmetric \( m \times m \) matrix that measures the degree of linear relationship within the data set between all possible pairs of variables (sensors). The subspaces in PCA are defined by the eigenvectors and eigenvalues of the covariance matrix as follow:

\[
\mathbf{C}_x \mathbf{P} = \mathbf{P} \Lambda,
\]

(2)

where the eigenvectors of \( \mathbf{C}_x \) are the columns of \( \mathbf{P} \), and the eigenvalues are the diagonal terms of \( \Lambda \) (the off-diagonal terms are zero). Columns of matrix \( \mathbf{P} \) are sorted according to the eigenvalues by descending order and they are called the Principal Components of the data set. The eigenvectors with highest eigenvalue represents the most important pattern in the data with the largest quantity of information. Choosing only a reduced number \( r \) of principal components, those corresponding to the first eigenvalues, the reduced transformation matrix could be imagined as a model for the structure. Geometrically, the transformed data matrix \( \mathbf{T} \) (score matrix) is the projection of the original data over the direction of the principal components \( \mathbf{P} \).

\[
\mathbf{T} = \mathbf{X} \mathbf{P}
\]

(3)

In the full dimension case, this projection is invertible (since \( \mathbf{P}^T \mathbf{P} = \mathbf{I} \)) and the original data can be recovered as \( \mathbf{X} = \mathbf{T} \mathbf{P}^T \). Now, with the given \( \mathbf{T} \), it is not possible to fully recover \( \mathbf{X} \), but \( \mathbf{T} \) can be projected back onto the original \( m \)-dimensional space and obtain another data matrix as follow:

\[
\tilde{\mathbf{X}} = \mathbf{P} \mathbf{T} = \mathbf{X} (\mathbf{P} \mathbf{P}^T)
\]

(4)

Considering \( \tilde{\mathbf{X}} \) as the projection of the data matrix \( \mathbf{X} \) onto the selected \( r \) principal components and \( \tilde{\mathbf{X}} \) as the projection onto the residual left components, the following decomposition can be performed:

\[
\mathbf{X} = \tilde{\mathbf{X}} + \tilde{\mathbf{X}}
\]

(5)

\[
\tilde{\mathbf{X}} = \mathbf{X} (\mathbf{P} \mathbf{P}^T)
\]

(6)

\[
\tilde{\mathbf{X}} = \mathbf{X} (\mathbf{I} - \mathbf{P} \mathbf{P}^T)
\]

(7)

### 3 Damage Detection and Localization

PCA applied as a pattern recognition is useful because permit to compare the state of a current structure with a baseline in order to determine if exist some changes and, besides whether these changes can be considered as a damage or not. To do this is necessary the use of different indices to localize and classify the different possible damages in the structure under test. The present section shown the mathematical formulation of some indices currently used by the author and some contributions methods based for damage localization.

#### 3.1 Damage indices for Detection

There are several kind of fault detection indices [2]. Two well-known indices are commonly used to this aim: the \( Q \)-index (or \( \text{SPE} \)-index), the Hotelling’s \( T^2 \)-statistic (\( D \)-statistic). The first one is based on analyzing the residual data matrix \( \tilde{\mathbf{X}} \) to represent the variability of the data projection in the residual subspace[9]. The second method is based in analyzing the score matrix \( \mathbf{T} \) to check the variability of the projected data in the new space of the principal components.

There exist another type of indices reported in the literature as combined index [13] and I index [4]. The first one is a combination of the \( Q \)-index and \( T^2 \)-index, the second one is used in meta-analysys and can be interpreted as a percentage of heterogeneity. In a general way, it is possible to define any index as it appears in the equation 8.

\[
\text{Index} = \mathbf{x}^T \mathbf{M} \mathbf{x}
\]

(8)

Where the vector \( \mathbf{x} \) represents measurements from all the sensors at a specific experiment trial, besides the matrix \( \mathbf{M} \) depends of the type of index as follows:

\[
\begin{align*}
\text{Q-index} &= \mathbf{x}^T \mathbf{M}_Q \mathbf{x} = \mathbf{x}^T (\mathbf{I} - \mathbf{P} \mathbf{T}) \mathbf{x} \\
\text{\( T^2 \)-index} &= \mathbf{x}^T \mathbf{M}_T \mathbf{x} = \mathbf{x}^T (\mathbf{P} \Lambda^{-1} \mathbf{P}^T) \mathbf{x} \\
\text{\( \varphi \)-index} &= \text{Q-index} + \text{T\textsuperscript{2}-index} = \mathbf{x}^T \mathbf{M}_x \mathbf{x} \\
\text{\( \phi \)-index} &= \mathbf{x}^T (\mathbf{I} - \mathbf{P} \mathbf{T} + \mathbf{P} \Lambda^{-1} \mathbf{P}^T) \mathbf{x} \\
\text{I-index} &= \mathbf{x}^T \mathbf{M}_I \mathbf{x}
\end{align*}
\]

(9) - (12)

where:

\[
\mathbf{M}_I = \begin{cases} 
0, & \text{for } Q \leq (k-1); \\
\frac{Q-(k-1)}{Q} \times 100\%, & \text{for } Q > (k-1).
\end{cases}
\]

(13)

#### 3.2 Contribution Methods for Localization

According to [2] five methods can be used for fault detection in process monitoring. Authors of this work pretends to adapt and use these methods for damage detection and localization in structures. These methodologies are used to calculate the contribution of each sensor to each index in each experiment trial. In this way, it is expected that the damage is located between actuator and sensor with largest contribution.

All the indices can determine whether there are damages and distinguish between them, however they does not provide reasons for it. The main idea is to determine which variable or variables are responsible. Variables with the largest contribution value are considered major contributors to the damage.

1) Complete Decomposition Contributions(CDC)

Complete decomposition Contributions also called contribution plots are well known diagnostic tools.
Decomposition Contribution (GDC) is defined as:

\[ Index = x^T M x = \| M^{\frac{1}{2}} x \|^2 \] (14)

\[ Index = \sum_{j=1}^{n} (\xi_j^T M^2 x)^2 = \sum_{j=1}^{n} CDC_{j}^{Index} \] (15)

\[ CDC_{j}^{Index} = x^T M^{\frac{1}{2}} \xi_j \xi_j^T M^2 x \] (16)

where \( \xi_j \) is the \( j \)th column of the identity matrix.

2) Partial Decomposition Contributions (PDC) This method decomposes a damage detection index as the summation of variable contributions.

\[ PDC_{j}^{Index} = x^T M \xi_j \xi_j^T x \] (17)

3) Diagonal Contributions (DC) The diagonal contribution remove the cross-talk among variables. The DC is defined as:

\[ DC_{j}^{Index} = x^T \xi_j \xi_j^T M \xi_j \xi_j^T x \] (18)

4) Reconstruction Based Contributions (RBC) The Reconstruction-Based Contribution [1] is an approach that uses the amount of reconstruction of a damage detection index along a variable direction as the contribution of that variable to the index. The RBC is defined as:

\[ RBC_{j}^{Index} = x^T M \xi_j (\xi_j^T M \xi_j)^{-1} \xi_j^T M x \] (19)

\[ RBC_{j}^{Index} = \frac{(\xi_j^T M x)^2}{(\xi_j^T M \xi_j)} \] (20)

5) Angle-Based Contributions (ABC)

\[ \xi_j = M^{\frac{1}{2}} \xi_j \] (21)

\[ \tau = M^{\frac{1}{2}} x \] (22)

The ABC of Variable \( j \) is the squared cosine of the angle between

\[ ABC_{j}^{Index} = \frac{(\xi_j^T x)^2}{\| \xi_j \|^2 \| x \|^2} = \frac{(\xi_j^T M x)^2}{\xi_j^T M \xi_j x^T M x} \] (23)

\[ ABC_{j}^{Index} = \frac{RBC_{j}^{Index}}{Index(x)} \] (24)

According to [2] it is possible to group these five methodologies in three general diagnosis methods. These are General Decompositional Contributions, Reconstruction Based Contributions, Diagonal Contributions. The complete and partial decomposition can be defined as special cases of this formulation. The General Decomposition Contribution (GDC) is defined as:

\[ GDC_{j}^{Index} = x^T M^{1-\beta} \xi_j \xi_j^T M^\beta x, \quad 0 \leq \beta \leq 1 \] (25)

When \( \beta = 0 \) \( PDC = GDC \) as is shown in the equations 26 - 29.

\[ GDC_{j}^{Index} = x^T M^{1-0} \xi_j \xi_j^T M^0 x \] (26)

\[ GDC_{j}^{Index} = x^T M^{1} \xi_j \xi_j^T (I) x \] (27)

\[ GDC_{j}^{Index} = x^T M^{1} \xi_j \xi_j^T x = PDC_{j}^{Index} \] (28)

When \( \beta = 1 \):

\[ GDC_{j}^{Index} = x^T M^{1-1} \xi_j \xi_j^T M^1 x, \quad IF \beta = 1 \] (29)

Here, \( I = \sum_{j=1}^{n} \xi_j^T \xi_j \), then is possible to reorganize the equation to obtain:

\[ GDC_{j}^{Index} = x^T M^{1-1} \xi_j \xi_j^T x = PDC_{j}^{Index} \] (30)

In the same way, when \( \beta = \frac{1}{2} \) is possible to obtain CDC.

\[ GDC_{j}^{Index} = x^T M^{1-\frac{1}{2}} \xi_j \xi_j^T M^{\frac{1}{2}} x \] (31)

\[ GDC_{j}^{Index} = CDC_{j}^{Index} \] (32)

Since ABC is a scaled version of RBC, it is possible to use RBC as a general case for both diagnosis methods.

4 Experimental Mockups

The approaches presented in this paper include the use of vibration time based signals and piezoelectric transducers to analyze the data provided from the structure. The analysis include the comparison between the signals collected from the healthy structure and the structure under test using statistical techniques. In general terms, the Structural Health Monitoring laboratory of the CoDAlab group utilized for doing the experiments contains:

- A chasis of National Instruments (NI-PXI 1033);
- A shelf to hang-up the elements to test.
- A card NI PXI-5412, it is a arbitrary waveform generator/oscilloscope of 250MS/s with 40mV to 40V input ranges.
- A chasis contain 5 slots, in each slot is possible to add cards of National instruments, for instance generation cards, acquisition cards, switches cards and other elements. The chasis can connect with a computer using the PXI port by an express card.
- A NI PXI-5114 card, it is a 8-bit Digitizer/oscilloscope with 14-bit resolution and 100 MS/s sampling rate.
- A Crosspoint Matrix Switch card. Using this card is possible to define until 4x32 matrix configuration. It is useful because depending of size of the structure and the number of sensors is possible to reconfigure the number of terminals to use.
- A laptop, to connect the chasis and to realize the programs of acquisition and processing data.
- A shelf to hang-up the elements to test.
Using the elements described above, the signals from the healthy and current structures are collected. The methodology applied in each experiment is:

1) The instrumented structure is suspend to isolate it from environment disturbances using elastic ropes.
2) One from all PZTs attached on the surface is chosen as the one which work as actuator by mean of the switch module.
3) An excitation signal is applied to the structure (vibrational input) with the chosen PZT and using the NI-generator card.
4) Vibrational responses at different points are recorded by using the rest of PZTs (sensors) and the digitizer card.
5) Actuator and sensors are changed using the switch module, and the steps 2 to 4 are repeated. These changes and repeats are automatically applied by the program developed in Labview.
6) Data in text based format is saved and organized.
7) Simulate damages or to do some real damages and repeat the steps 2 until 6.
8) Apply the strategy based on PCA to compare the vibrational responses of the current and healthy structures.

The second structure is a smooth-raw aluminium 2015-3501 (figure 3). The dimensions of this plate is $25cm \times 25cm \times 0.1cm$. This plate was instrumented with four piezoelectric transducer discs (PZT’s) attached on the surface as is shown in the figure 3. A real damage was made between the PZT2 and PZT4, 300 experiments were performed and recorded: 100 using the undamaged structure, and 200 using the structure with different size of damage (increasing the depth). The PCA model was created using 100 % of the whole dataset of the undamaged structure, 50 % of the dataset of undamaged structure and the whole dataset of the damaged structure were used for testing the approach.

**5 PRELIMINARY RESULTS**

**5.1 Damage Detection**

The damage detection is a methodology previously reported in [10], and uses the scores and the indices $T^2$ and $Q$ (figure 4), for this, one PCA model is built in each phase (PZT1 as actuator, PZT2 as actuator, and so on) using the signals recorded by sensors during the experiments with the undamaged structure. Data from the experiments using the current structure (damaged or not) are projected on the model. Projections onto the first principal components (scores), $T^2$-statistics and $Q$-statistics are calculated by each PCA model.

Each PCA model by itself is used as tool for damage, for this, the scores and $T$ vs $Q$ plots are performed for
Fig. 4. Damage detection methodology

each PCA model and the analysis is developed. Figures 5 and 6 show the scatter plots of score 1 vs score 2 and $T^2$ vs $Q$, respectively for the PZT 1 in the aircraft turbine blade.

As seen in the figures 5 and 6, some damages are clearly distinguished from undamaged structure, also it is possible to distinguish damages between them in both plots.

The same methodology has been applied using other indices (see figures 7, 8).

Fig. 5. Scores with the PZT1 as actuator

Fig. 6. $T^2$ vs $Q$ plot with the PZT1

5.2 Damage Classification

The damage classification methodology (see figure 9) was previously reported by Mujica et al. [10] to analyze one phase, now authors extended this methodology in 2010 [3] to include the combination of all the phases in order to use the information collected of each PZT.

The goal of this work is to organize, to combine and to contrast the information obtained from all models in order to provide a general diagnosis of the structure. To do that, any classification tool can be used, authors have chosen Self Organizing Maps (SOM). This Artificial Neural Network (ANN) is an unsupervised algorithm known also as Kohonen Map [6]. Inputs to the SOM are the projections onto the new space, $T^2$–index and $Q$–index of each phase.
5.3 Damage Localization

The damage localization methodology used in this work has been previously suggested in [10], this methodology only include the analysis for one phase and one actuator, currently the authors of this paper extended this methodology to make a combined analysis of all phases and all sensors [11].

In the methodology (see fig. 11), one PCA model is built in each phase as was done in the damage detection and the projections onto the first principal components (scores), $T^2$-index, $Q$-index, $\phi$-index and $I$-index are calculated by each PCA model.

To localize the damage, the contribution of each sensor in the different phases to each index in each phase is calculated and finally are accumulated in order to obtain one measure that show the region with more abnormalities, this contribution is a measure of the level of influence of the damage to the sensor. In each phase, a region of the structure is selected as the region where the damage is located. Considering all phases, a general diagnosis could be performed (intersection of all the regions).

• Aircraft Turbine Blade

Figures 12, 13, 14 and 15 show the contribution of each PZT (1 to 7) to each index ($T^2$, $Q$, $\phi$ and $I$) at a specific trial (experiment using the structure to diagnose). In this case, the structure has the damage 3 (D3 is a damage located near to PZT 4, see fig. 17).

Analyzing Figure 12 (contributions to $Q$-index) it is possible to observe the following: During the phase 1 (PZT1 as actuator), the highest contribution is obtained in the PZT4. Therefore, the damage is located between PZT1 and PZT4. A similar situation is founded in phases 2, 3, 5, 6 and 7; so, the damage is located between PZT2 and PZT4 (phase 2), PZT3 and PZT4 (phase 3), PZT5 and PZT4 (phase 5), PZT6 and PZT4 (phase 6), and PZT7 and PZT4 (phase 7).
In a general way, the region or area of the damage can be defined as the intersection of the different areas found in each phase. In this case, it can be concluded that the damage is nearby to PZT4.

![Fig. 13. Contributions of each PZT to $T^2$ – index](image)

![Fig. 14. Contributions of each PZT to $\varphi$ – index](image)

![Fig. 15. Contributions of each PZT to $I$ – index](image)

From Figure 13 (contributions to $T^2$ – index) it can be seen that some contributions are negative. Negative contributions do not have any physical sense and therefore, they are not considered in the analysis. Results are similar to obtained by using $Q$ – index. The main difference appears in phase 6, the highest contribution is obtained by PZT7. Contributions to $\varphi$ – index (Figure 14) and to $I$ – index (Figure 15) show similar results. The highest contribution is carry out by PZT4.

In order to show the final diagnosis (considering contributions at all phases) it is necessary to specify areas in the structure that consider paths between actuator and sensors. The contribution of each sensor in each phase defines the weight of the path (region between actuator and sensor). Finally, the sum of all the weighted regions establishes the region where the damage is located.

![Fig. 16. Interface of Localization of damage](image)

![Fig. 17. Localization of the damage 3 using contributions to $Q$ – index](image)

From Figure 16 it can be seen the software application developed in Matlab. Here, the image of the structure is loaded, and the position of every PZT is manually defined (using the mouse). A typical path-planning algorithm is implemented to define the areas between PZT’s. The final diagnostic of the current structure (with damage 3) using contributions to $Q$ – index presented in Figure 17 (the higher the value of the color, the more probability of the localization of the damage). As it is expected, the damage is located near to PZT4. Additionally, performing experiments with damage 1 present in the structure, the approach identifies the localization of the damage near to PZT1 as is shown in Figure 13.

- Aluminium plate
The five methods explained in concepts section were applied to each index and compared using the damage located between PZT 2 and PZT 4 (fig. 19) with different depth, figures 20-24 shows the results of using the DC, CDC, PDC, ABC and RBC for the $Q^{-index}$ and damage 1, the higher the value of the color, the more probability of the localization of the damage. As it is expected, the damage is located between PZT2 and PZT4.

Figure 31 shows the contributions obtained for each path between the different PZT’s for the five methods. As shown, in each method, the path between PZT 2- PZT 4 contains the highest values of contribution, this is because the damage is located in this path. With all the methods is possible to locate the damage with different values, the lowest contribution is found with CDC method, additionally, it is possible to see that in all the methods the difference
Fig. 24. Damage localization: using RBC

Fig. 25. Damage 1: comparison of methods

Fig. 26. Damage 2 using diagonal contributions

Fig. 27. Damage 2 using complete decomposition contributions

Fig. 28. Damage 2 using partial decomposition contributions

Fig. 29. Damage 2 using angle based contributions

between the values of each path for each method are significantly greater for the path between PTZ 2 and PZT 4 except for CDC method where the values are similar.

The same five methodologies are applied for the $T^2$-index in the same damage with different depth, results are shown in figures 26-30, the higher the value of the color, the more probability of the localization of the damage. As it is expected, the damage is located between PZT2 and PZT4. Figure 15 shows the contributions obtained for each path between the different PZT’s for the five methods. As for the $Q$-index, in each method, the path between PZT 2- PZT 4 contains the highest values of contribution, this is because the damage is in this path. In this case compared with the results for $Q$-index exist more
Fig. 30. damage 2 using reconstruction based contributions

Fig. 31. damage 2: comparison of methods

differences between the different paths using CDC method.

6 CONCLUSION

The application of PCA to detect, localize and classify damages in structures were presented. Each application include the use of a methodology which include an active piezoelectric system, PCA and some statistic indices.

A novelty multiactuator piezoelectric system for localization of damages has been developed. The approach combines strategies to study: (i) The dynamic or vibrational response of the structure at different exciting and receiving points. (ii) The correlation of these dynamical responses when some damage is presented in the structure by using PCA and some statistical measures that can be used as indices of damage. (iii) The influence of every sensor in the indices, this contribution can be used to localize the origin of the change in the vibrational characteristic (damage).

The approach proposed for classifying were tested with excellent results using PCA analysis (first two scores, $T^2$, $Q$, $\phi$ and $I−I$-index). The methodology developed to localize damages is based in PCA models built from vibrational responses of the structure using a active system [3]. The region which contain the damage is obtained by finding the highest value area, for this, the sum of the contributions obtained for each sensor to each index is calculated.

ACKNOWLEDGMENT

This work has been supported by the “Ministerio de Ciencia e Innovación” in Spain through the coordinated research project DPI2008-06564-C02-01/02. The authors would like to thank the support from the “Agència de Gestió d’Ajuts Universitaris i de Recerca” of the “Generalitat de Catalunya”, “Escola Universitària d’Enginyeria Tècnica Industrial de Barcelona (EUETIB)”, and “Universitat Politècnica de Catalunya”.

REFERENCES