Damage Detection Using Robust Fuzzy Principal Component Analysis

F. GHARIBNEZHAD, L. E. MUJICA, J. RODELLAR
and C.-P. FRITZEN

ABSTRACT

In this work Robust Fuzzy Principal Component Analysis (RFPCA) is used and compared with comparing with classical Principal Component Analysis (PCA) to detect and classify damages. It has been proved that the RFPCA method achieves better result mainly because it is more compressible than classical PCA and also carries more information, hence not only it can distinguish the healthy structure from the damaged structure much sharper than the traditional counterparts but also in some cases traditional PCA is incapable of discerning the pristine from damaged structure. This work involves experimental results using pipe-like structure powered by a piezoelectric actuators and sensors.

INTRODUCTION

Structural health monitoring (SHM) has gained a significant amount of attention in the research and industrial communities over the last two decades. The concept of actively monitoring structures for damage is of interest because it presents the ability to detect and locate damage in a structure before it can propagate and cause serious failure. Damage can be defined as changes introduced into a system that adversely affects its current or future performance [1]. The ability to know when and where damage has occurred in a structure can reduce the costs associated with scheduled inspections and the repair of failed structures, and also improve the overall safety of the structure. SHM can be applied to many sectors of infrastructure including civil, mechanical, and aerospace systems.

[1,2,3] Department of Applied Mathematics III. Universitat Politécnica de Catalunya. Comte d'Urgell 187, 08036 Barcelona, Spain; [4] Institute of Mechatronics and Control Engineering – Mechanics, University of Siegen, Germany;
Email: fahit.gharibnezhad@upc.edu, luis.eduardo.mujica@upc.edu, jose.rodellar@upc.edu, fritzen@imr.mb.uni-siegen.de

Several damage detection methods have been developed and used in this field. Among them, feature discrimination using Principal Component Analysis (PCA), has received a significant consideration due to unique features that this method offers [2, 3, 4]. PCA is a popular statistical method which tries to explain the covariance structure of data by means of a small number of components. These small quantity of components are calculated based on maximizing variance and decomposing covariance. Usually, two or three PCs provide a good summery of all the original variables. Indeed, PCA follows two most significant goals; Firstly, it reduces the dimension of data. Secondly, it can also reveal those underlying factors or combinations of the original variables that principally determine the structure of the data distribution [5].

Despite the interesting characteristic that PCA carries, it is suffering from some flaws such as sensitivity to outliers, missing data and poor linear correlation between variables due to poorly distributed variables [6]. Therefore, data reduction, modeling, and any other method based on classical PCA become unreliable if the mentioned drawbacks are not considered [7].

To overcome these limitations, some methods have been proposed such as the method which is based on the eigenvectors of robust covariance matrix [8] and the method which is based on projection pursuit (PP) [9, 10] and also the method based on both approach which attempts to combine the advantages of both methods [11].

The superiority of mentioned methods over classical PCA in SHM for distinguishing damages has been considered before by the authors of this paper [12].

Another most illuminating approach is to use robust fuzzy principal component analysis (RFPCA).It has been proved that RFPCA method achieves better result mainly because it is more compressible than classical PCA, i.e. the first fuzzy principal components accounts for significantly more of the variance than their classical counterparts. Therefore, by carrying more information in primary PCs, it can provide more information for any damage detection approach based on it and as a result, distinguishing much better between the healthy and damaged structures [6].

To support the claims mentioned above, this work involves an experimental benchmark with pipe-like structure equipped with piezoelectric transducers as actuators and sensors. Damages have been simulated by saw cuts in different severity. Then, the ability of separating non-damaged structure from damages with different severity has been compared between classical PCA and RFPCA and it has been shown that RFPCA is more efficient to distinguish between undamaged, real damages.

The remaining parts of this paper are organized as follows. First, a theoretical consideration on traditional and robust fuzzy PCA is explained. After that, experimental setup is described. Furthermore, results and discussion are presented and discussed. Finally, conclusions are drawn.

THEORITICAL CONSIDERATION

Traditional PCA

Principal Component Analysis (PCA) is an important and essential technique for data reduction, image compression and feature extraction [13]. PCA is a tradeoff between clarity of representation and ease of understanding. The main motivation of

PCA is to project the data from a high dimensional space onto a lower dimensional space and reveal the part of information that is not easy distinguished in original data. To do this PCA transforms the original variable into new, uncorrelated variables called principal components (PC) which are linear combinations of the original variables and generally demonstrate the data more feasible in much less dimension. However, it is well-known that PCA, as with any other multivariate statistical method is sensitive to outliers, missing data and poor linear correlation between variables [14]. The reader is referred to [15, 16] for more information.

Robust Fuzzy Principal Component Analysis

To alleviate the drawbacks of traditional PCA, different methods have been suggested [17, 5, 6, 11]. The ability of them to provide more feasible results in SHM has been considered elsewhere [12, 7]. The Robust Fuzzy Principal Component analysis learning rule that is used in this work is described in [13] and it is based on the approach proposed in [18].

Considering a data set with n observations like $X = \{x_1, \dots, x_n\}$ the optimization function E subject to $u_i \in \{0,1\}$ is defined by:

$$E(U, w) = \sum_{i=1}^{n} u_i e(x_i) + \eta \sum_{i=1}^{n} (1 - u_i). \tag{1}$$

The goal is to minimize E with respect to U and w, where $U = \{u_i, i = 1, \dots, n\}$ is the membership sets and η is the threshold. Since u_i is the binary variable and, w is the continuous variable, the optimization with gradient descent approach is hard to solve using gradient descent. Therefore, a new objective function is proposed by [13] as follows:

$$E = \sum_{i=1}^{n} {u_i}^{m_1} e(x_i) + \eta \sum_{i=1}^{n} (1 - u_i)^{m_1} \tag{2}$$

Subject to $u_i \in [0,1]$ and $m_1 \in [1, \infty)$. Now u_i being the membership of x_i belonging to data cluster and $(1 - u_i)$ is the membership of x_i belonging to noise cluster. m_1 is the so-called fuzziness variable. In this case, $e(x_i)$ measures the error between x_i and the class center. This idea is similar to the C-means algorithm [19].

Since now u_i is a continuous variable, the difficulty of mixture of discrete and continuous optimization can be avoided and a gradient descent approach can be used. Firstly, the gradient of equation (2) is computed respect to u_i and equaled to zero, therefore:

$$u_i = \frac{1}{1+(e(x_i)/\eta)^{1/(m_1-1)}}\,. \tag{3}$$

By substituting this membership in equation (2) the following equation is obtained:

$$E = \sum_{i=1}^{n} \left(\frac{1}{1+(e(x_i)/\eta)^{1/(m_1-1)}}\right)^{m_1-1} e(x_i) \tag{4}$$

On the other hand, the gradient respect to w is

$$\frac{\partial E}{\partial w} = \beta(x_i)\left(\frac{\partial e(x_i)}{\partial w}\right), \tag{5}$$

where,

$$\beta(x_i) = \left(\frac{1}{1+(e(x_i)/\eta)^{\frac{1}{m_1-1}}}\right)^{m_1}, \tag{6}$$

and m_1 is the fuzziness variable. If $m_1 = 1$, the fuzzy membership reduces to the hard membership and can be determined by following rule:

$$u_i = \begin{cases} 1 & if(e(x_i)) < \eta \\ 0 & otherwise \end{cases} \tag{7}$$

Now η is a hard threshold in this situation. There is no general rule for the setting of m_1, but most papers set $m_1 = 2$. In [13], authors derived the following algorithm for the optimization procedure

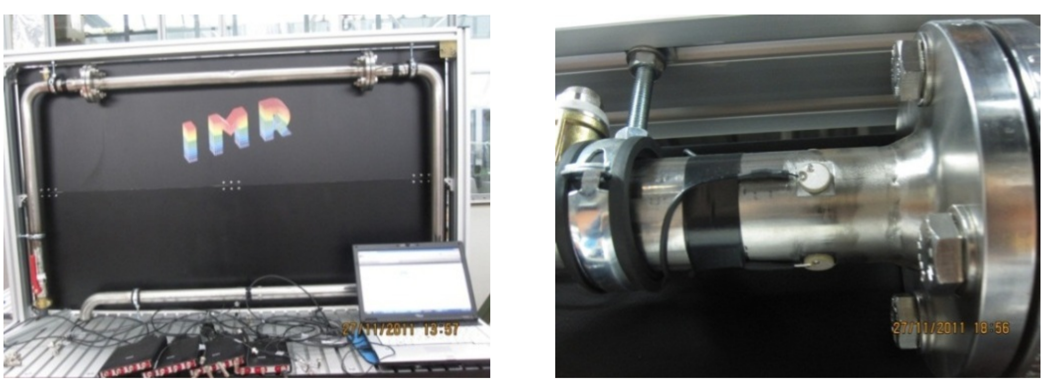
1. Initially set the iteration count $t = 1$, the iteration bound T, learning coefficient $\alpha_0 \in (0,1]$, soft threshold η to a small positive value and randomly initialize the weight w.
2. While t is less than T, do steps 3–9.
3. Compute $\alpha_t = \alpha_0(1 - \frac{t}{T})$, set $i = 1$ and $\sigma = 0$
4. For the number of observations , do steps 5–8
5. Compute y=$w^T x_i, u = yw$ and $v = w^T u$
6. Update the weight as : $w^{new} = w^{old} + \alpha_t \beta(x_i)[y(x_i - u] + (y - v)x_i]$
7. Update the temporary count $\delta = \delta + e_1(x_i)$
8. Add 1 to i.
9. Compute $\eta = (\delta/n)$ and add 1 to t.

The weight w in the updating rules converges to the principal component vector almost surely [20, 21].

EXPERIMENTAL SETUP

The RFPCA algorithm mentioned above is applied on a data set captured from pipe benchmark consist of damages (cuts) with different severities. Damage detection in pipe like structures has gotten strong consideration due to their importance in industry as many works have been done to detect flaws in pipes and tubes using guided waves [22, 23, 24].

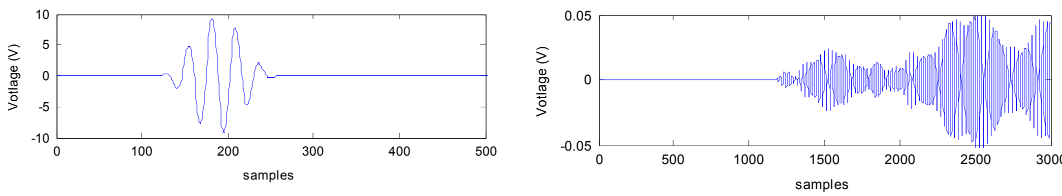
Figure ***1*** shows the mentioned testing powered by a data acquisition system, Handyscope® that allows generating and capturing signals using the mounted PZT actuator/sensors.



(a) (b)

Figure 1. Pipe-like benchmark a) structure and acquisition system b) mounted transducers.

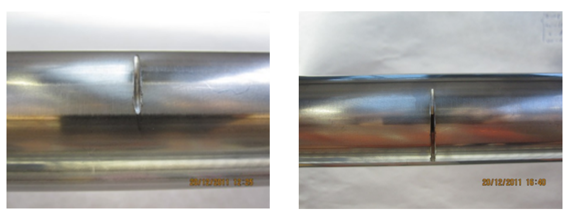
Waves are generated by consecutive by actuating the 4 PZTs on the left side of the pipe using a 180 KHz tone burst signal with the amplitude of 10 volts and these are captured by each of the 4 PZTs on the right side. The PZTs are numbered in order of up, rear, down and front in both sides starting from actuators (i.e. PZT 1 correspond to left side and upper position and so on). Therefore, 16 routes exists each route connecting an actuator on the left side and a sensor on the right side. Each experiment is repeated 20 times for each route. To reduce noise, a mean signal is saved to obtain one observation. Finally, 30 observations are captured for each route in each state (different damages and pristine structure). *Figure 2* shows a sample of the generated and received waves after a decimating by 10 and denoising procedure.



(a) (b)

Figure 2. Generated and received waves, a) tone burst actuating signal b) received signal.

Damages are presented by adding measured cuts with different depths in order to simulate damages with different severities (0.75 mm crack depth, 2 mm crack depth and 16 mm width, 2 mm crack depth and 30 mm wide , complete hole) *Figure 3* shows an example of mentioned damages. Damages are placed in a line between actuator 2 and sensor 6.



(a) (b)

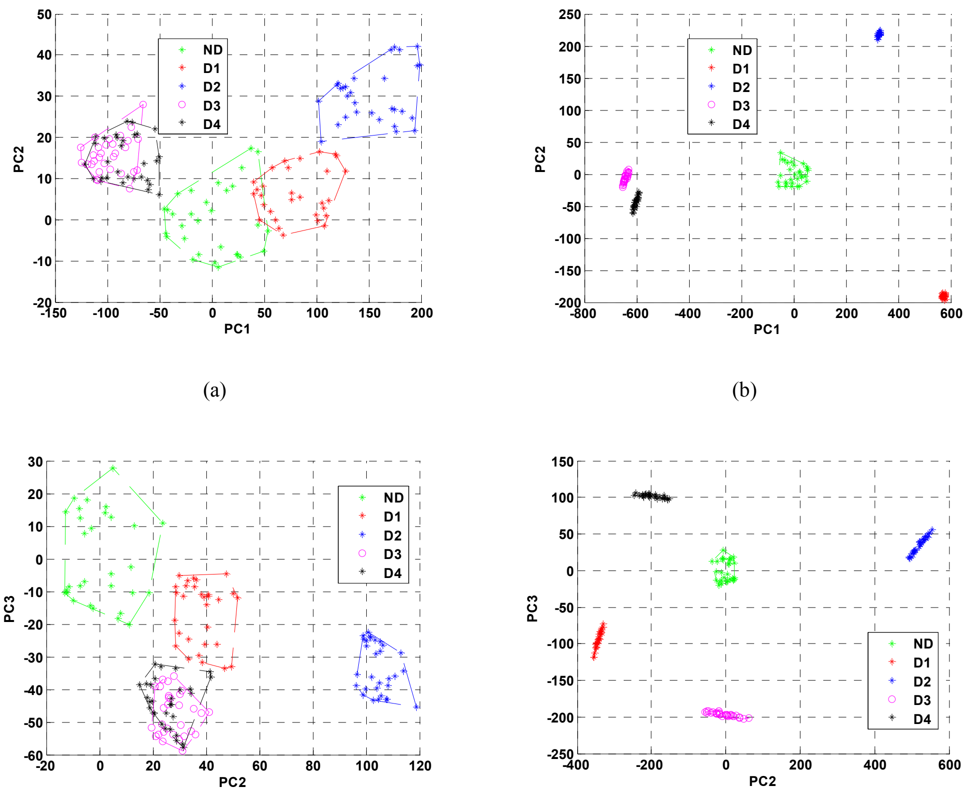
Figure 3. (a) and (b) two example of saw cuts in pipe structure with different severities.

RESULTS AND DISCUSSION

To build the baseline model of the healthy structure, traditional PCA and RFPCA are applied on the data matrix that contains dynamical responses of structure in form of $\mathbf{X}_{n\times m}$, where n represents the number of observations, and m the number of variables (samples). In a traditional approach [2], the projection matrix $\mathbf{P}$ is calculated which consists of eigenvectors of the covariance matrix of baseline data. The columns of matrix $\mathbf{P}$ are sorted according to the eigenvalues by descending order, whereas the eigenvector with the highest eigenvalue represents the most important pattern in the data with the largest quantity of information. This matrix is used as a model to apply the testing data, which contain data from both: the damaged and pristine structure. 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 $\mathbf{T}_{n\times r}$ (score matrix) is the projection of the original data onto the direction of the principal components **P** as follows:

$$\mathbf{T} = \mathbf{XP} \tag{7}$$

The projection is a representation of data points in the principal component space. As PCA derives the best possible r dimensional ($r < p$) representation of the Euclidean distances among objects, new points can be used directly to show the similarity or difference between observations in a much lower dimension space. The same philosophy is applied to RFPCA, where iteration weight w is considered as a transformation matrix as described in the algorithm. For instance, *Figure 4* shows the score depicted in the first-second and second-third PCs space both in PCA and RFPCA.

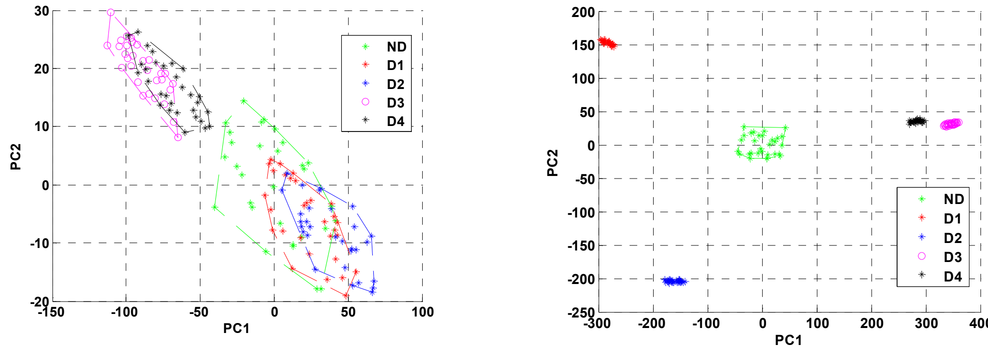


(a) (b)

(c) (d)

Figure 4. Traditional PCA vs RFPCA a) scores on first and second PCs using PCA b) scores on first and second PCs using RFPCA c) second and third PCs using PCA d) second and third PCs using RFPCA.

From figure 4 it can be seen, in this route (from actuator 1 to sensor 6) that classical PCA can distinguish between damaged and healthy but it is unable to distinguish all damages from each other (i.e. damage 3 and 4 have overlap), whereas RFPCA provides much better separation between all the patterns related to any state of structure. It should be mentioned that in some routes, classical PCA even is not able to separate the healthy and damaged structure with minimum damage severity state, while FPCA can achieve this goal in all routes as depicted in *Figure 5*.



(a) (b)

Figure 5. Route 1-6 from actuator 1 to sensor 6 a) PCA cannot separate the pattern of damages from the healthy one b) complete separation is satisfied.

The main reason for achieving better results by RFPCA is that it is more compressible than classical PCA, i.e. the primary principal components count for significantly more of the variance than their classical counterparts hence, they convey more information from the structure and as thereupon better result is achieved.

CONCLUSION

A Robust Fuzzy Principal Component Analysis (RFPCA) method has been applied in this work to detect and classify damages in a pipe-like structure. The efficiency of the new approach is proved by its ability to provide much sharper differentiation of the patterns rather than its classical counterpart. It has been proved that damage detection based on robust PCA is more reliable as in all cases sharp separation of patterns are achieved whereas the traditional one suffers to distinguish the pattern clearly. More work is expected to show the superiority of robust method when other algorithms for damage detection are applied based on principal component analysis.

Acknowledgment

This work has been supported by the "*Ministerio de Economía y Competitividad*" in Spain through the Coordinated Project DPI2011-28033-C03-01, and the "*Formación de Personal Investigador*" FPI doctoral scholarship. The author would like to thank greatly the support from "*Institut für Mechanik und Regelungstechnik , Universitat Siegen*" and all of its members especially *Dr.-Ing. Maksim Klinkov* and *Dr.-Ing. Peter Kraemer* (for the test bed), *Mr. Gerhard Dietrich* (help with the experiment) and Msc. *Ms. Inka Buethe* and Msc. *Mr. Miguel Angel Torres-Arredondo* (for helping to collect the data).

REFERENCES

[1] Sohn, H., Farrar, C. R., Hemez, F. M., Czarnecki, J. J., Shunk, D. D., Stinemates, D. W. & Nadler, B. R. 2003 A Review of Structural Health Monitoring Literature: 1996–2001. Los Alamos National Laboratory Report, LA-13976-MS.

[2] L. E. Mujica, J. Rodellar, A. Fernandez, and A. Guemes, “Q-statistic and T2-statistic PCA-based measures for damage assessment in structures,” *Structural Health Monitoring*, Nov. 2010.
[3] P. De Boe and J.-C. Golinval, “Principal Component Analysis of a Piezosensor Array for Damage Localization,” *Structural Health Monitoring*, vol. 2, no. 2, pp. 137–144, Jun. 2003.
[4] Fan, Lili , C. U. of Hong Kong. Dept. of Building and Construction, *Structural health monitoring based on principal components analysis implemented on a distributed and open system*. City University of Hong Kong, 2006.
[5] H. Pop, “Principal Components Analysis based on a fuzzy sets approach,” *Mij*, vol. 1, no. 2, p. 1, 2001.
[6] C. Sarbu, “Principal component analysis versus fuzzy principal component analysis:: A case study: the quality of danube water (1985-1996),” *Talanta*, vol. 65, pp. 1215–1220, 2005.
[7] F. Gharibnezhad and L. M. Delgado, “Damage detection in the presence of outliers based on robust PCA,” in *Proceedings of th 8th International Conference on Structural Dynamics, EURODYN 2011*, Leuven: 2011, 2011.
[8] C. Croux, P. Filzmoser, and M. Oliveira, “Algorithms for Projectionâ€“Pursuit robust principal component analysis,” *Chemometrics and Intelligent Laboratory Systems*, vol. 87, no. 2, pp. 218–225, Jun. 2007.
[9] C. Croux and a. Ruizgazen, “High breakdown estimators for principal components: the projection-pursuit approach revisited,” *Journal of Multivariate Analysis*, vol. 95, no. 1, pp. 206–226, Jul. 2005.
[10] M. Hubert, “A fast method for robust principal components with applications to chemometrics,” *Chemometrics and Intelligent Laboratory Systems*, vol. 60, no. 1-2, pp. 101–111, Jan. 2002.
[11] M. Hubert, P. Rousseeuw, and K. V. Branden, “ROBPCA: a new approach to robust principal component analysis,” *Technometrics*, pp. 1–34, 2005.
[12] F. Gharibnezhad, L. E. Mujica, and J. Rodellar, “Comparison of two robust PCA methods for damage detection in presence of outliers,” *Journal of Physics: Conference Series*, vol. 305, no. 1, p. 12009, 2011.
[13] T.-N. Yang and S.-D. Wang, “Robust algorithms for principal component analysis,” *Pattern Recognition Letters*, vol. 20, no. 9, pp. 927–933, Sep. 1999.
[14] T. Cundari, C. Sârbu, and H. Pop, “Robust Fuzzy Principal Component Analysis (FPCA). A Comparative Study Concerning Interaction of Carbon-Hydrogen Bonds with Molybdenum-Oxo Bonds,” *Journal of chemical information and computer sciences*, vol. 42, no. 6, pp. 1363–1369, 2002.
[15] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2002, vol. 2.
[16] J. Shlens, “A tutorial on principal component analysis,” *Systems Neurobiology Laboratory, University of California at San Diego*, 2005.
[17] H. F. Pop and C. Sârbu, “A new fuzzy regression algorithm.” *Analytical chemistry*, vol. 68, no. 5, pp. 771–8, Mar. 1996.
[18] L. Xu and A. Yuille, “Robust principal component analysis by self-organizing rules based on statistical physics approach,” *Neural Networks, IEEE Transactions on DOI - 10.1109/72.363442*, vol. 6, no. 1, pp. 131–143, 1995.
[19] P. Luukka, “Classification based on fuzzy robust PCA algorithms and similarity classifier,” *Expert Systems with Applications*, vol. 36, no. 4, pp. 7463–7468, May 2009.
[20] E. Oja, “A simplified neuron model as a principal component analyzer,” *Journal of mathematical biology*, vol. 15, no. 3, pp. 267–73, Jan. 1982.
[21] E. Oja and J. Karhunen, “On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix,” *Journal of mathematical analysis and applications*, vol. 106, no. 1, pp. 69–84, 1985.
[22] H. Shin, “Guided wave tuning principles for defect detection in tubing,” *Journal of nondestructive evaluation*, vol. 17, no. 1, pp. 27–36, 1998.
[23] J. L. ROSE, D. JIAO, and J. J. SPANNER, “engUltrasonic guided wave NDE for piping,” *engMaterials evaluation*, vol. 54, no. 11, pp. 1310–1313.
[24] J. L. Rose, J. J. Ditri, A. Pilarski, K. Rajana, and F. Carr, “A guided wave inspection technique for nuclear steam generator tubing,” *NDT & E International*, vol. 27, no. 6, pp. 307–310, Dec. 1994.