Experimental validation of Automated structural health monitoring based on analysis datas

Zheng He, Xiaoyu Sun *, Xuan Gu, Qingbo Sui, Ju Liu, Kuo Yuan

College of Aerospace and Civil Engineering, Harbin Engineering University, Harbin 150001 China

*sunxiaoyu520634@163.com

Abstract

Structural health monitoring aims to detect damages in man-made engineering structures by monitoring changes in their vibration response. Unsupervised learning algorithms can be used to obtain a model of the undamaged condition and detect which new samples of the structure are not in agreement with it. However, in real structures with a sensor network configuration, the number of candidate features usually becomes large. Therefore, complexity increases and it is necessary to perform feature selection and/or dimensionality reduction to achieve good detection accuracy. In this paper, we propose to exploit the three-way structure of data and apply a true multi-way data analysis algorithm: Parallel Factor Analysis. A simple model is obtained and used to train novelty detectors. The methods are tested both with real and simulated structural data to assess that the three-way analysis can be successfully used in structural health monitoring. Furthermore, the usefulness of the approach for feature selection is also analyzed.

Keywords

Novelty detection; Multi-way analysis; Damage detection; Structural health monitoring; Sensor network; Vibration monitoring; Feature selection; Data analysis.

1. Introduction

The purpose of *structural health monitoring* (SHM) is to identify damages in engineering structures [1]. To achieve that, it compares certain damage-sensitive features, extracted from measured data of the monitored structure, to those of a baseline model of the normal condition. Therefore, when the model is obtained from data, this process can be framed as a problem of intelligent feature extraction and machine learning [2]. This is the usual case in real large structures where, generally, limited prior knowledge about the model is available.

A three-way decomposition is suitable because data in this application are naturally represented as a *three-way structure* with J features measured in K different locations during I time periods. This type of data has been successfully analyzed in other application fields, such as multichannel electroencephalograms (EEG) [3] or chemometrics [4], by means of multi-way data analysis algorithms[5-6], which take advantage of the additional modes of data to produce simple and robust models. *Parallel Factor Analysis*(PARAFAC) has been selected among those methods because it presents certain advantages regarding simplicity [4]. Indeed, the aim is to produce parsimonious models that are useful in understanding the monitored structure and result in better damage detection. This approach is tested on three different structures and the results are assessed with four standard novelty detection algorithms. They are evaluated in terms of the area under the ROC (receiver operating characteristic) curve [7], which gives a single number measure of the decision threshold-independent diagnostic accuracy of the classifiers.

2. Experimental

Parallel Factor Analysis, also known as *Canonical Decomposition*, is a method for decomposing Nth-order tensors that extend the idea of bilinear factor models to multilinear data [4]. It was originally proposed in[8-9]. For the three-way case, this algorithm decomposes data into triads or trilinear components. The result is given by three loading matrices, A, B and C, with elements *aif*, *bjf* and *ckf*, where f=1,…,F and *F* is the number of components. A graphical illustration can be seen in Fig. 1.

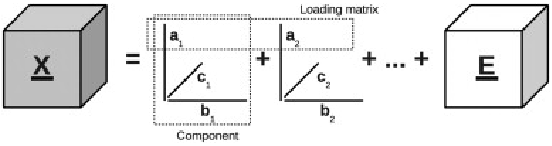

Fig. 1. Parallel Factor Analysis decomposition.

The first experiment uses data from a wooden bridge structure in laboratory conditions [10]. This structure was investigated also in our previous work [11-12]. An electrodynamic shaker excites its vertical, transverse and torsional modes to simulate random ambient excitation in the structure. Small masses are attached to simulate damages in seven different locations. The heaviest ones only weigh 0.5% of the mass of the structure, which is 36 kg. The bridge is instrumented with 15 sensors, which measure acceleration in two directions, as shown in Fig. 2, and whose locations are not optimized according to the known locations of the damages. Data were acquired with a sampling frequency 256 Hz. We considered transmissibility magnitudes below 120 Hz, obtained by means of a Fast Fourier Transform (FFT) [13]. Among 105 possible pairs of sensors, only those linking adjacent ones were considered. This choice was suggested by the concept of nearest-neighbor coupling [14], which assumes that transmissibility between two certain sensors is only sensitive to the changes between them when a single excitation is applied anywhere except near a boundary condition. The undamaged measurements are randomly distributed between the training and the test sets for each iteration: 70% of them are used for training whereas the rest is used for the test set, along with all the faulty samples.

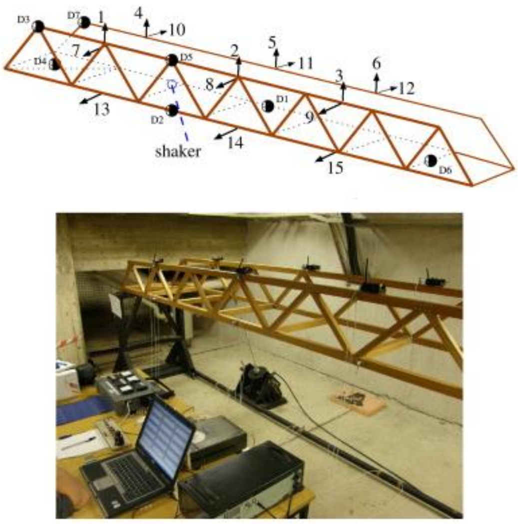

Fig. 2. Bridge structure and placement of the sensors.

Fig. 3 shows the damage detection results for the three experiments. From them all, a common conclusion can be drawn: Both parameters are important to achieve an optimal diagnostic accuracy. Indeed, the results suggest that larger averaging windows produce more accurate detectors. The noisy nature of the signals used to train the detectors explains this finding, which confirms the results pointed out in [15]. High variance and the presence of outliers in data make it difficult to obtain good classifiers, so it seems advisable to use the largest window, as long as there are enough data to train a good detector. From a wireless sensor network (WSN) perspective, this is a good characteristic, since it contributes to reduce transmission rate. Wireless sensor nodes are usually equipped with a limited power source and data communication consumes a significant amount of energy, so sensor lifetime is strongly affected by the transmission power efficiency. Nevertheless, if on-line monitoring is considered, it will also be necessary to take into account how often the SHM reports are expected.

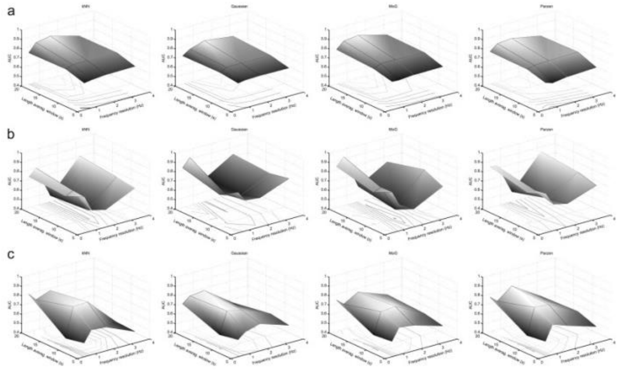

Fig. 3. AUROCs for varying averaging window lengths and frequency resolutions with standard PARAFAC. (a) Bridge. (b) Bookshelf. (c) Beam.

It was explained before that the *number of components* is a free parameter of the PARAFAC algorithm. For that reason, the behavior of the novelty detection methods was analyzed with respect to the number of factors, using the optimal combination of resolution and averaging according to the previous results. From the results shown in Fig. 4, we can conclude that the number of components selected by CORCONDIA, which only considers undamaged data and is done before novelty detection, is optimal for the bridge and bookshelf structures and slightly underestimated for the simulated beam.

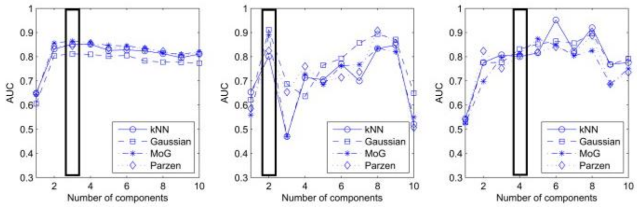

Fig. 4. Performance with respect to the number of components for bridge, bookshelf and beam. The highlighted number is the one selected by CORCONDIA for each structure.

3. Results and discussion

It can be seen that, in general, the distribution of the instances labeled as damaged is different to the baseline model. Nevertheless, there are some undistinguishable points that generally correspond to damages caused by the lighter weights, which may not change the response of the bridge enough. This is shown in Fig. 5, although this figure must be interpreted with care, because PARAFAC scores are not expressed on an orthonormal basis [4]. However, detectability varies for different weights and positions of the damage. A line plot of the components reveals that damages tend to be strongly detectable in one of the components whereas they show values close to normal in the other ones, and those discriminant components are different depending on damage location (see Fig. 6). For instance, component 1 shows a mismatch between the model and the instances corresponding to damage 5, whereas component 2 shows a clear mismatch for damages 4 and 7 (these damages have been annotated in Fig. 2 for a clear interpretation). It should be noted that the first component of the spatial mode gives the highest score to the pair 7–8 and the second component to the pair 1–4 (see Fig. 7). These locations highlighted in the decomposition of the training data are consistent with the situation of the corresponding damages.

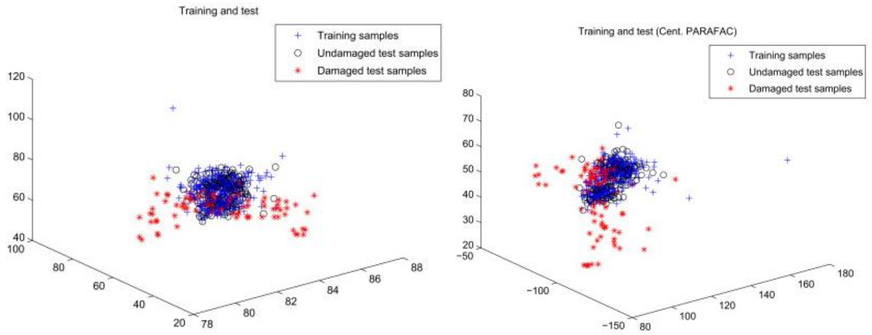

Fig. 5. Training and test data of the bridge experiment for plain PARAFAC and PARAFAC on centered data. Damaged and undamaged test cases are labeled with different markers.

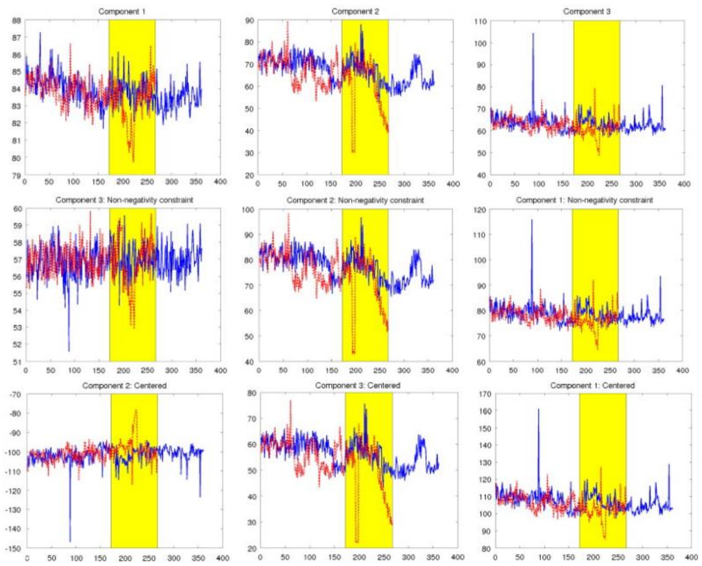

Fig. 6. Visual comparison of training (blue solid line) and test data (red dashed line) of the bridge experiment in the three components. The shaded area is used to represent the test instances labeled as damaged. In the first row, the results obtained with PARAFAC are shown. The second row shows the results nonnegative PARAFAC and the third row the ones for PARAFAC on centered data. The components have been rearranged to make them match and ease the visualization. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

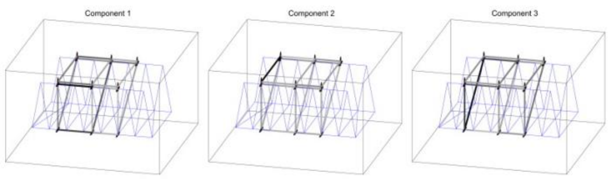

Fig. 7. The plots show the sensor pair scores for each component in the bridge experiment. Connections with a dark color represent higher scores.

For the bookshelf structure, the input data were obtained from a 4096-point FFT and with averaging windows of 20.48 s. The number of components for the decompositions was set to 2 using the CORCONDIA criterion (30×640×12). The analysis of the results shows, in general, similar effects to the ones discussed before. However, in this case, centering apparently succeeds in projecting data onto a space where the decomposition is computed more easily. Again, detection of damage is easier in certain components, as the one shown in Fig. 8. Nevertheless, there is no interpretable association in this experiment between components with high damage detectability and scores of loading matrices. Nearest neighbors classifier gives the worst detection results.

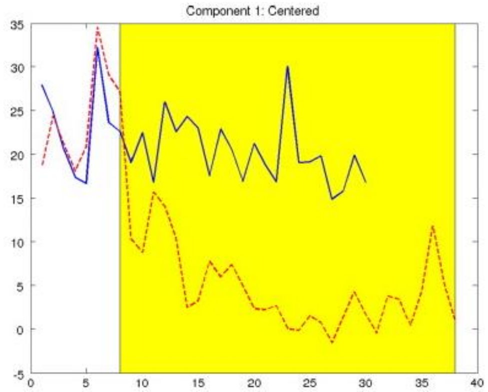

Fig. 8. Abnormal behavior in damaged instances of the bookshelf. Obtained with PARAFAC on centered data.

Four components were used for the cantilever beam. Features were obtained every 20 s by averaging the ones obtained with a 2500-point FFT (20×160×19). Detectability of damages is again high in this experiment. In this case, both the regular decomposition and the one on centered data provide better results than the constrained decomposition. One component shows a visible drift from the baseline model (Fig. 9) in every decomposition. That agrees with the growing nature of damage. Decompositions of training data always show a high score for the pair 9–10, located in the middle of the structure, in the discriminant component. The best results in this experiment are achieved by the k-NN classifier.

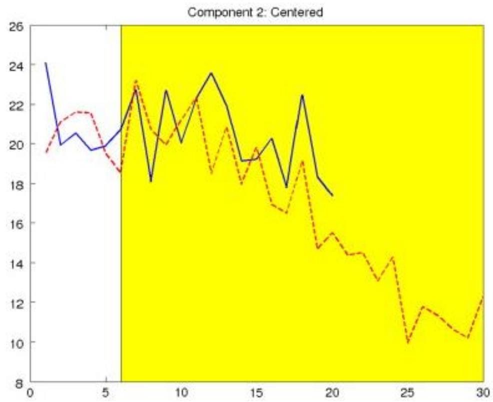

Fig. 9. Drift caused by increasing damage in the cantilever beam. Obtained with PARAFAC on centered data.

4. Conclution

We presented a structural health monitoring problem as a novelty detection problem in the three-way analysis setting. The three aspects of data are time windows, frequencies and sensor pairs. We demonstrated that this approach can yield good damage detection results as well as a criterion to select the features of the system.

Acknowledgements

This work is supported by the National Natural Science Foundation of China（No. 11602066） and the National Science Foundation of Heilongjiang Province of China (QC2015058 and 42400621-1-15047), the Fundamental Research Funds for the Central Universities.

References

[1] C.R. Farrar, K. Worden An introduction to structural health monitoring Philos. Trans. R. Soc. A, 365 (1851) (2007), pp. 303–305

[2] K. Worden, G. Manson The application of machine learning to structural health monitoring Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.: Math., 365 (1851) (2007), pp. 515–537
[3] W. Deburchgraeve, P. Cherian, M.D. Vos, R. Swarte, J. Blok, G. Visser, P. Govaert, S.V. Huffel Neonatal seizure localization using PARAFAC decomposition Clin. Neurophysiol., 120 (10) (2009), pp. 1787–1796
[4] A.K. Smilde, R. Bro, P. Geladi Multi-Way Analysis with Applications in the Chemical Sciences John Wiley and Sons (2004)
[5] E. Acar, B. Yener Unsupervised multiway data analysis: a literature survey IEEE Trans. Knowl. Data Eng., 21 (1) (2009), pp. 6–20
[6] M. Mørup Applications of tensor (multiway array) factorizations and decompositions in data mining Wiley Interdiscip. Rev.: Data Mining Knowl. Discovery, 1 (1) (2011), pp. 24–40
[7] A.P. Bradley The use of the area under the ROC curve in the evaluation of machine learning algorithms Pattern Recognition, 30 (7) (1997), pp. 1145–1159
[8] R.A. Harshman Foundations of the PARAFAC procedure: models and conditions for an “explanatory” multi-modal factor analysis UCLA Work. Pap. Phonetics, 16 (1970), pp. 1–84
[9] J.D. Carroll, J.J. Chang Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart–Young decomposition Psychometrika, 35 (1970), pp. 283–319
[10] J. Kullaa Elimination of environmental influences from damage-sensitive features in a structural health monitoring system D.L. Balageas (Ed.), Proceedings of the First European Workshop on Structural Health Monitoring 2002, DEStech Publications Inc., Onera (2002), pp. 742–749
[11] J. Toivola, J. Hollmén Feature extraction and selection from vibration measurements for structural health monitoring Advances in Intelligent Data Analysis VIII, IDA 2009, Lecture Notes in Computer Science, vol. 5772, Springer (2009), pp. 213–224
[12] J. Toivola, M.A. Prada, J. Hollmén Novelty detection in projected spaces for structural health monitoring P.R. Cohen, N.M. Adams, M.R. Berthold (Eds.), Advances in Intelligent Data Analysis IX, Lecture Notes in Computer Science, IDA 2010, vol. 6065, Springer (2010), pp. 208–219
[13] S.K. Mitra Digital Signal Processing: A Computer-Based Approach (second ed.)McGraw-Hill, New York, NY, USA (2001)
[14] T.J. Johnson, D.E. Adams Transmissibility as a differential indicator of structural damage J. Vib. Acoust., 124 (4) (2002), pp. 634–641
[15] K. Worden, G. Manson, D. Allman Experimental validation of structural health monitoring methodology: part I. Novelty detection on a laboratory structure J. Sound Vib., 259 (2) (2003), pp. 323–343