Applicability of an output-only structural damage detection based on transmissibility measurements and kernel principal component analysis

A. Santos, M. F. M. Silva, R. Santos, E. Figueiredo, N. M. M. Maia and J. C. W. A. Costa

***Abstract*— Frequency response functions have been employed as damage-sensitive features in the vibration-based structural damage detection. The need for measuring the excitation forces arises as a remarkable limitation on the application of those features in real-world applications. Transmissibility measurements can be explored as features with output-only nature, which imply the need for measuring only the response signals. Thus, an output-only damage detection method is proposed, combining transmissibilities with kernel principal component analysis. The results performed on transmissibility measurements from a laboratory steel beam reveals that the output-only method has high potential to be applied in a wide range of monitoring solutions.**

***Keywords*— Structural health monitoring, Transmissibility measurements, Damage detection, Kernel principal component analysis.**

I. INTRODUCTION

In the structural health monitoring (SHM) field, damage detection based on the vibration response measurements from engineering structures has become a crucial research area due to its potential to be applied in real-world scenarios [1]. From the vibration signals, damage-sensitive features can be extracted and used to assess early and progressive structural damage via appropriate data treatment.

Frequency response functions (FRFs) play an important role in the vibration-based damage detection area [2]. Many works have been employed the FRFs as features to generate some kind of damage indicators (DIs) that reveal the structural condition of monitored structures [3], [4]. However, the need for measuring the excitation forces arises as a remarkable limitation on the application of FRFs in real-world SHM solutions. As an alternative, transmissibility measurements have been widely explored as features in damage detection for SHM [5], [6], [7], due to their output-only nature, i.e. the need for measuring only the response signals.

Considering the successful use of transmissibilities to distinguish between undamaged and damaged conditions of monitored structures, instead of generating a DI directly from the transmissibility measurements, several studies have been attempting to combine FRFs and transmissibilities with machine learning (ML) algorithms to detect structural damage.

The combination between FRFs or transmissibilities and artificial neural networks (ANN) has been widely investigated. In a first manner, the transmissibilities are the input of a complex ANN that should detect damage simulated as stiffness changes in a structure [8], [9]. However, the number of spectral lines of the transmissibilities defines the number of input nodes of the ANN, which may lead to a complex neural architecture and expensive computational load. In other attempt, the FRFs, computed from a monitored railway wheel, have their dimensionality reduced via principal component analysis (PCA) and this reduced form is the input of the ANN [10]. Although acceptable results were achieved with this technique, a drawback appears when the transmissibilities from undamaged and damaged cases must be known in advance by PCA.

An approach that uses outlier analysis, density estimation and auto-associative neural network combined with measured transmissibilities was proposed to assess damage in aerospace structures [11]. In this case, some parts of the transmissibility measurements are selected as features in a visual manner, which may not be generalized to other health monitoring cases.

More recently, FRFs and transmissibilities were linked to the approach based on Mahalanobis squared distance (MSD) to determine the structural condition of a monitored beam via different types of DIs [12]. Similar to a first proposal [13], the MSD algorithm presents problems, namely numerical errors to compute a large covariance matrix, when all spectral lines of the damage-sensitive features are considered. Note that the ML algorithms, such as MSD and ANN, often work with a large number of observations (or measurements) and a small dimensionality (or spectral lines).

PCA [14] and its nonlinear version, kernel PCA (KPCA), have been applied to reduce the dimensionality of the original features such that a trade-off should be reached in the sense that the appropriate dimensionality needs to be not only large enough to account for all normal condition but also small enough to be as sensitive as possible to damage. To reach this appropriate reduced form of transmissibilities, this paper proposes an output-only structural damage detection method in the context of a statistical pattern recognition (SPR) paradigm. In this paradigm, the proposed method is divided into two phases [15]: feature extraction and feature classification. In the first phase, the dimensionality of the transmissibilities

A. Santos is with Faculty of Computing and Electrical Engineering, F. 17, Q. 4, L. E., Marabá, 68505–080, Pará, Universidade Federal do Sul e Sudeste do Pará.

M. F. M. Silva, R. Santos, and J. C. W. A. Costa are with Applied Electromagnetism Laboratory, Federal University of Pará, R. Augusto Corrêa 01, Belém, Pará, 66075110 Brazil.

E. Figueiredo is with Faculty of Engineering, Universidade Lusófona de Humanidades e Tecnologias, Campo Grande 376, Lisbon, 1749-024 Portugal.

N. M. M. Maia is with LAETA, IDMEC, Instituto Superior Técnico, Departamento de Engenharia Mecânica, Universidade de Lisboa, Lisbon, 1049-001 Portugal.

is reduced by application of KPCA. In the second phase, based on outlier detection, DIs are generated for new transmissibilities considering the data model computed in the first phase. The proposed method is validated on a monitored steel beam, where the damage classification performance is evaluated based on false-positive indications of damage (Type I errors) and false-negative ones (Type II errors).

The remainder of this study is organized as follows. In section II, a brief background related to transmissibility measurements is emphasized and the feature extraction and feature classification phases of the proposed method are described. A description of the test structure is provided in section III, along with a summary of the measurement sets. Experimental results on measurement sets are discussed in section IV; a comparison with the PCA algorithm is also presented. Finally, section V highlights a discussion related to the strengths and challenges of the proposed method.

II. Output-only method

This section deals with the background and phases related to the method proposed in this study.

A. Transmissibility measurements

Transmissibility measurements are defined as relations between motion responses and motion reference-responses [16]. In practice, it is often convenient to obtain the transmissibilities without the knowledge of the excitation forces. In particular, the direct or scalar transmissibility measurement $T_{i,j}\left(w\right)$ between an output i and and reference-output j is defined as the ratio between the two response spectra,

$$T_{i,j}\left(w\right)=\frac{Y_i\left(w\right)}{Y_j\left(w\right)}, \tag{1}$$

where $Y_i\left(w\right)$ and $Y_j\left(w\right)$ are the complex amplitudes of the responses $y_i\left(t\right)$ and $y_j\left(t\right)$, respectively, for a harmonic force applied at a given coordinate, and for each frequency w.

In general, for a random input, a transmissibility measurement can be estimated in several manners. The most common option is using an output-only H_1 estimator by dividing an estimate of the cross-power spectrum $S_{i,j}\left(w\right)$ between the output $Y_i\left(w\right)$ and the reference-output $Y_j\left(w\right)$ by an estimate of the auto-power spectrum $S_{j,j}\left(w\right)$ from the reference-output $Y_j\left(w\right)$,

$$T_{i,j}\left(w\right)=\frac{S_{i,j}\left(w\right)}{S_{j,j}\left(w\right)}. \tag{2}$$

A set of transmissibilities is acquired by measuring responses on all coordinates and directions of interest at the structure divided by the reference response from the same fixed measurement coordinate. One should consider a training set, $\mathbf{X}\in\mathbb{R}^{n\times d}$, with d-dimensional transmissibilities from n different conditions when the structure is undamaged and a test set, $\mathbf{Z}\in\mathbb{R}^{l\times d}$, where l is the number of transmissibilities from the undamaged and/or damaged conditions.

B. Feature extraction

For the feature extraction phase, the KPCA algorithm [17] is used to reduce the dimensionality of the transmissibility measurements. Let $\mathcal{X}\in\mathbb{R}^d$ be the input space such that the transmissibilities $\mathbf{x}_i\in\mathcal{X}$, $i=1,\ldots,n$. Every transmissibility $\mathbf{x}$ is then mapped to a $d_{\boldsymbol{\phi}}$-dimensional feature space $\mathcal{H}$ by applying the mapping functions ϕ_m, $m=1,\ldots,d_{\boldsymbol{\phi}}$, where

$$\boldsymbol{\phi}(\mathbf{x})=\begin{bmatrix}\phi_1(\mathbf{x}) & \phi_2(\mathbf{x}) & \ldots & \phi_{d_{\boldsymbol{\phi}}}(\mathbf{x})\end{bmatrix}^\top. \tag{3}$$

By employing the kernel trick [18], $\mathcal{K}:\mathcal{X}\times\mathcal{X}\mapsto\mathbb{R}$ is defined as a positive semi-definite scalar kernel function satisfying for all $\mathbf{x}_i$, $\mathbf{x}_j\in\mathcal{X}$,

$$\mathcal{K}\left(\mathbf{x}_i,\mathbf{x}_j\right)=\boldsymbol{\phi}\left(\mathbf{x}_i\right)^\top\boldsymbol{\phi}\left(\mathbf{x}_j\right). \tag{4}$$

$\mathcal{K}\left(\cdot\right)$ defines an inner product that allows to map the transmissibilities implicitly to a high-dimensional kernel space. Let

$$\boldsymbol{\Phi}=\left[\boldsymbol{\phi}(\mathbf{x}_1)\quad\boldsymbol{\phi}(\mathbf{x}_2)\quad\ldots\quad\boldsymbol{\phi}(\mathbf{x}_n)\right] \tag{5}$$

be the $d_{\boldsymbol{\phi}}\times n$ matrix of the mapped transmissibilities and $\mathbf{K}=\boldsymbol{\Phi}^\top\boldsymbol{\Phi}$ be the $n\times n$ kernel (Gram) matrix. According to Mercer's theorem, any continuous, symmetric, and positive semi-definite function that maps $\left(\mathbf{x}_i,\mathbf{x}_j\right)$ onto a high-dimensional feature space can represent a kernel [19]. The kernel trick then consists of specifying the kernel $\mathcal{K}\left(\cdot\right)$ instead of the mapping $\boldsymbol{\phi}$. Herein, a Gaussian kernel [20] is employed,

$$\mathcal{K}\left(\mathbf{x}_i,\mathbf{x}_j\right)=\exp\left(-\frac{\parallel\mathbf{x}_i-\mathbf{x}_j\parallel^2}{2\sigma^2}\right), \tag{6}$$

where this kernel implicitly defines a high-dimensional feature space with a bandwidth σ^2.

To avoid that the first principal component becomes much larger than the other components, the kernel matrix $\mathbf{K}$ should be replaced by a centered version [17],

$$\mathbf{K}\rightarrow\mathbf{K}-\frac{\mathbf{1_n}}{n}\mathbf{K}-\mathbf{K}\frac{\mathbf{1_n}}{n}+\frac{\mathbf{1_n}}{n}\mathbf{K}\frac{\mathbf{1_n}}{n}, \tag{7}$$

with $\mathbf{1_n}$ as the $n\times n$ matrix composed of ones.

The eigenvalues $\boldsymbol{\Sigma}$ and the corresponding eigenvectors $\mathbf{U}$ can be then derived by using singular value decomposition (SVD) to solve the generalized eigenvalue problem [17],

$$\mathbf{KU}=\mathbf{U\Sigma}. \tag{8}$$

Afterwards, the $\boldsymbol{\Sigma}_1$ and $\mathbf{U}_1$ should be defined as follows,

$$\boldsymbol{\Sigma}=\left[\boldsymbol{\Sigma}_1\quad\boldsymbol{\Sigma}_2\right],\boldsymbol{\Sigma}_1\in\mathbb{R}^{r\times r};\mathbf{U}=\left[\mathbf{U}_1\quad\mathbf{U}_2\right],\mathbf{U}_1\in\mathbb{R}^{n\times r}, \tag{9}$$

where $\boldsymbol{\Sigma}_1$ comprises the r largest eigenvalues and $\mathbf{U}_1$ the corresponding eigenvectors. The transmissibility measurements $\mathbf{X}$ can be then represented in a reduced form as the transpose of

$$\mathbf{X}_p=\sqrt{\boldsymbol{\Sigma}_1}\mathbf{U}_1^\top. \tag{10}$$

In addition, the estimation of parameters required for the KPCA algorithm is discussed in the following. There are multiple methods to optimize the bandwidth parameter σ^2 of the Gaussian kernel [21]. However, those methods require that $n \geq d$ and when transmissibility measurements are used, often $n < d$. Thereby, a rule-of-thumb is employed,

$$\mathbf{v} = \mathrm{var}\left(\mathbf{X}\right), \sigma^2 = \frac{n}{d}\sum_{i=1}^{d} v_i, \tag{11}$$

where $\mathrm{var}\left(\mathbf{X}\right)$ is the variance of the training data. Several criteria have been proposed to determine the number of principal components r retained in the high-dimensional feature space [22]. In this study, r is derived to comprise nearly all normal variability of the training data,

$$\sqrt{\frac{\sum_{i=1}^{r} \mathbf{\Sigma}_{i,i}}{\sum_{i=1}^{n} \mathbf{\Sigma}_{i,i}}} \geq 0.99, \tag{12}$$

being $\mathbf{\Sigma}_{i,i}$ the i-th diagonal element of $\mathbf{\Sigma}$. Note that the variance retained in the standard PCA is usually 0.9–0.95 [22]. On the other hand, the KPCA has been often used with 0.99 because there are potentially n nonzero principal components for the Gaussian kernel.

C. Feature classification

Since an undamaged data model was established in the previous phase by training the KPCA, in this phase a DI is generated to any new transmissibility measurement $\mathbf{z}_i \in \mathbb{R}^d$, $i = 1, \ldots, l$.

First, a new transmissibility measurement should be mapped onto the high-dimensional feature space in the form of $\mathbf{\Phi}\left(\mathbf{z}_i\right)^\top \mathbf{\Phi}$ (or $\mathbf{\Phi}^\top\mathbf{\Phi}\left(\mathbf{z}_i\right)$), by using $\mathbf{X}$ and $\mathbf{z}_i$ in Equation 6. Besides, a centering should be performed, such as,

$$\mathbf{\Phi}\left(\mathbf{z}_i\right)^\top \mathbf{\Phi} \rightarrow \mathbf{\Phi}\left(\mathbf{z}_i\right)^\top \mathbf{\Phi} - \frac{\check{\mathbf{1}}_\mathbf{n}}{n}\mathbf{K} - \mathbf{\Phi}\left(\mathbf{z}_i\right)^\top \mathbf{\Phi}\frac{\mathbf{1}_\mathbf{n}}{n} + \frac{\check{\mathbf{1}}_\mathbf{n}}{n}\mathbf{K}\frac{\mathbf{1}_\mathbf{n}}{n}, \tag{13}$$

with $\check{\mathbf{1}}_\mathbf{n}$ as the $l \times n$ matrix where all elements are equal to 1.

Second, the eigenvectors $\mathbf{U}_1$ should be replaced by a normalized version,

$$\mathbf{u}_m \rightarrow \frac{\mathbf{u}_m}{\sqrt{\mathbf{\Sigma}_{m,m}}}, m = 1, \ldots, r, \tag{14}$$

being the transmissibilities used in the test phase, $\mathbf{Z}$, represented in a reduced form as

$$\mathbf{Z}_p = \mathbf{\Phi}\left(\mathbf{z}_i\right)^\top \mathbf{\Phi}\mathbf{U}_1. \tag{15}$$

Finally, a DI is generated for the i-th new transmissibility measurement as follows,

$$\mathrm{DI}\left(\mathbf{z}_i\right) = \mathbf{\Phi}\left(\mathbf{z}_i\right)^\top \mathbf{\Phi}\mathbf{U}_1\mathbf{U}_1^\top\mathbf{\Phi}^\top\mathbf{\Phi}\left(\mathbf{z}_i\right). \tag{16}$$

III. Test structure and data sets

For the experimental validation of the proposed method, a steel beam was used, with rectangular cross-section, dimensions $1002\times35\times6$ mm^3 and weight 1.740 kg. Two inextensible cables simulating "free-free" support conditions suspended the test structure. The test structure was excited at location 3 with a pseudo-random signal by a Brüel & Kjaer 4809 shaker, powered by a Brüel & Kjaer 2706 power amplifier. The force was transmitted through a stinger and measured by a Brüel & Kjaer 8200 force transducer; the responses were measured by 23 piezoelectric CCLD accelerometers (equally spaced coordinates). The response signals were fed into the multichannel data acquisition unit Brüel & Kjaer 2816 (PULSE) and analyzed with the Labshop 6.1 Pulse software installed on the attached laptop. The experimental setup is shown in Figure 1 and more details can be found in Sampaio *et al* [4].

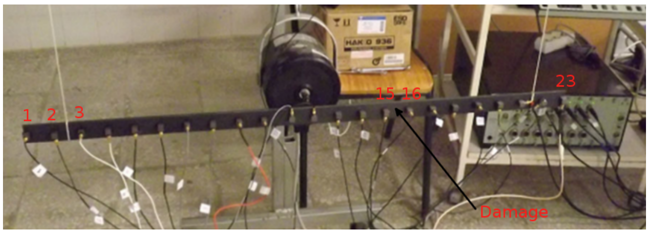

Fig. 1. Experimental setup with the identification of the accelerometers and damage location.

The damage was simulated with a reduction in the height by a saw cut. Basically, saw cuts, with several depths, were inflicted to the beam between locations 15 and 16 to create nine damage levels, as synthesized in Table I.

TABLE I
DAMAGE LEVELS INFLICTED TO THE EXPERIMENTAL BEAM BY SAW CUTS.

Damage level	Width (mm)	Depth (mm)
D01	1.0	0.5
D02	1.0	1.0
D03	1.0	1.25
D04	1.0	1.6
D05	1.0	3
D06	1.0	3.5
D07	1.0	4
D08	1.0	4.5
D09	1.0	5

Therefore, from Table I, one can infer that the response signals from the accelerometers deployed on the beam were measured in 10 conditions; the undamaged or baseline one (D00), and the nine levels of damage (D01 to D09) inflicted in the middle of the locations 15 and 16. Furthermore, the frequency range used for the analysis of the beam was 0-800 Hz (3200 spectral lines) and 15 averages have been taken to acquire the accelerations.

Assuming a force at location 3, 30 measurement sets have been performed from the instrumented beam. The first measurement set is the undamaged beam (Baseline condition – BC). The next two measurement sets are also of the undamaged beam. The following 27 measurement sets correspond to the nine saw cuts of the beam with three measurement sets each (Damaged condition – DC).

Considering all possible combinations to output and reference-output, 506 transmissibilities can be generated for each measurement set. For convenience, only 100 transmissibilities are selected at random from each measurement set, yielding 3000 transmissibilities to compose the training and test measurement sets, where the first 300 are transmissibilities from the undamaged condition and the last 2700 transmissibilities corresponding to the nine levels of damage. The training data is composed of 90% of the transmissibility measurements from the undamaged condition. The remaining 10% of the transmissibility measurements are used during the test phase to make sure that the DIs do not fire off before the damage starts and to evaluate the level of generalization of the proposed method. The test data is composed of all the measurement sets, even the ones used during the training phase. Note that the training process is unsupervised, which imposes serious limitations for a cross-validation procedure.

IV. Results and Discussion

Examples of transmissibilities and FRFs derived from the monitored beam under different conditions are shown in Figure 2. There are notable differences between the undamaged condition and the maximum level of damage for both transmissibilities and FRFs. However, the large number of spectral lines and small number of measurements make the transmissibilities inappropriate to be suitably processed by the ML algorithms. Thereby, this fact highlights the need for dimensionality reduction.

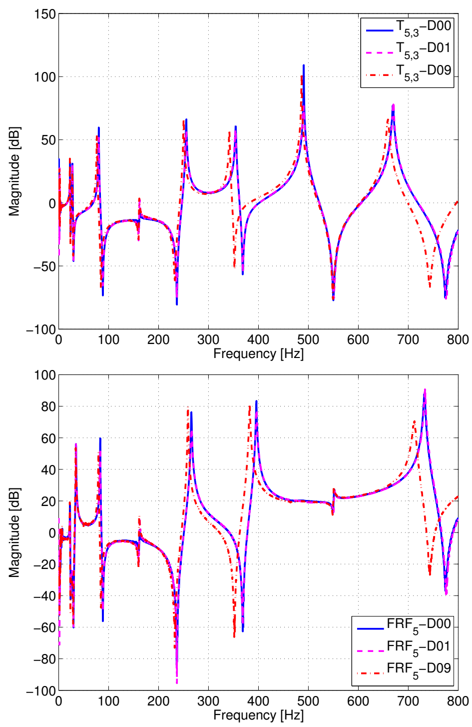

Fig. 2. $T_{5,3}\,(w)$ (top) and $\mathrm{FRF}_5\,(w)$ (bottom) for three different conditions.

After the algorithms have reduced the dimensionality of the transmissibility measurements, the DIs derived from the KPCA and PCA algorithms are shown in Figure 3, along with a threshold defined for a level of significance of 5% over the training data. The KPCA, which selected 181 principal components, can maintain a monotonic relationship between the progressive level of damage and the amplitude of the DI and minimizes quite well the Type I and II errors. In opposition, the PCA, which selected only 14 principal components, fails to achieve the monotonic relationship and exhibits many false-negative indications of damage. The performances of the algorithms are summarized in Table II, where approximately 13 Type I errors are expected due to the threshold selected for 95% of confidence in the training data and the disparity between the total errors from both algorithms is evident.

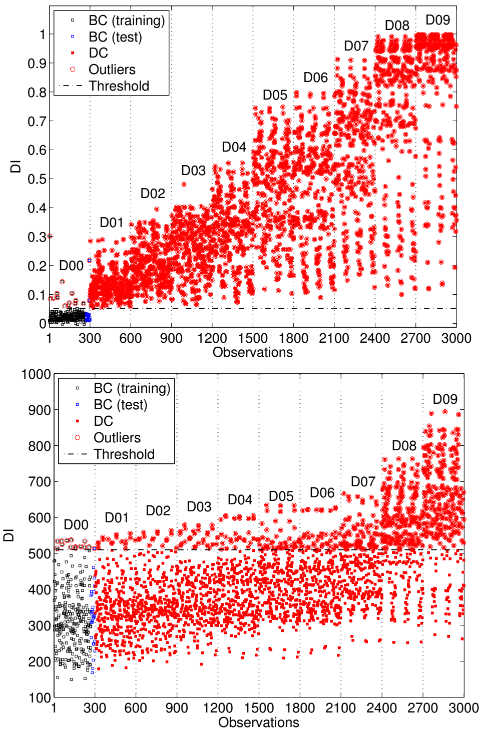

Fig. 3. Outlier detection based on the KPCA (top) and PCA (bottom) algorithms.

TABLE II

Number and percentage of Type I and Type II errors.

Algorithm	Error		
	Type I	Type II	Total
KPCA	15 (5.00%)	1 (0.037%)	16 (0.53%)
PCA	14 (4.67%)	2032 (75.26%)	2046 (68.20%)

The large difference between the performances of the algorithms is explained through Figure 4. In the high-dimensional feature space mapped by KPCA, the principal components are distributed in a more representative manner than those distributed onto the feature space projected by PCA. Thus, the transmissibilities that have been reduced, in the training phase, from 3200 to 181 dimensions can be generalized to new transmissibilities, ensuring an adequate dimension to detect structural anomalies, whereas the reduction from 3200 to 14 dimensions by PCA may impact in an underfitting regarding the training measurement sets.

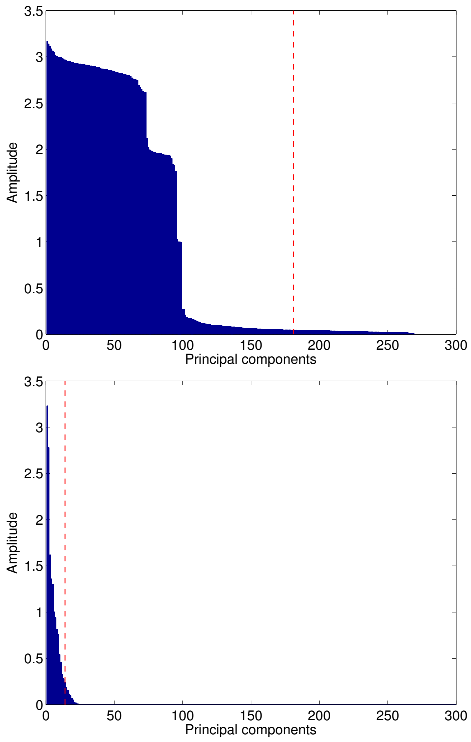

Fig. 4. Distribution of principal components for the KPCA (top) and PCA (bottom) algorithms. The number of principal components selected is highlighted by a red dashed line.

V. Conclusions

This paper proposed an output-only structural damage detection method, for which transmissibility measurements and ML algorithms were used to assess the condition of monitored structures. Feature extraction and feature classification phases were developed, based on the SPR paradigm, to reduce the dimensionality of transmissibilities via KPCA (training) and to generate a DI that establishes the level of damage for each new transmissibility measurement (test), respectively.

The damage detection performances on the test scenario confirmed that the proposed method is better than the alternative one. When the KPCA was compared to the PCA, the improvement of the mapped feature space proved to have a direct and positive impact on the dimensionality reduction step and consequently damage detection. This explains, in part, the relatively poor performance of the PCA on measurement sets from the test experiment.

Unlike the other approaches, the output-only strategy introduced in this study can process the transmissibility measurements by KPCA such that an appropriate dimensionality is achieved for ML applications, solving the problem of many spectral lines and small number of measurements.

Acknowledgements

This work was supported in part by the National Council for the Improvement of Higher Education and the National Council for Scientific and Technological Development, Brazil.

References

[1] W. Fan and P. Qiao, "Vibration-based damage identification methods: a review and comparative study," *Structural Health Monitoring*, vol. 10, no. 1, pp. 83–111, 2011.

[2] N. Maia, J. Silva, E. Almas, and R. Sampaio, "Damage detection in structures: from mode shape to frequency response function methods," *Mechanical Systems and Signal Processing*, vol. 17, no. 3, pp. 489–498, 2003.

[3] R. Sampaio and N. Maia, "Strategies for an efficient indicator of structural damage," *Mechanical Systems and Signal Processing*, vol. 23, no. 6, pp. 1855–1869, 2009, special Issue: Inverse Problems.

[4] R. Sampaio, N. Maia, R. Almeida, and A. Urgueira, "A simple damage detection indicator using operational deflection shapes," *Mechanical Systems and Signal Processing*, vol. 72–73, pp. 629–641, 2016.

[5] C. Devriendt and P. Guillaume, "The use of transmissibility measurements in output-only modal analysis," *Mechanical Systems and Signal Processing*, vol. 21, no. 7, pp. 2689–2696, 2007.

[6] N. M. Maia, R. A. Almeida, A. P. Urgueira, and R. P. Sampaio, "Damage detection and quantification using transmissibility," *Mechanical Systems and Signal Processing*, vol. 25, no. 7, pp. 2475–2483, 2011.

[7] Y.-L. Zhou, E. Figueiredo, N. Maia, and R. Perera, "Damage Detection and Quantification Using Transmissibility Coherence Analysis," *Shock and Vibration*, vol. 2015, pp. 1–16, 2015.

[8] K. Worden, "Structural fault detection using a novelty measure," *Journal of Sound and Vibration*, vol. 201, no. 1, pp. 85–101, 1997.

[9] Q. Chen, Y. Chan, and K. Worden, "Structural fault diagnosis and isolation using neural networks based on response-only data," *Computers & Structures*, vol. 81, no. 22–23, pp. 2165–2172, 2003.

[10] C. Zang and M. Imregun, "Structural damage detection using artificial neural networks and measured FRF data reduced via principal component projection," *Journal of Sound and Vibration*, vol. 242, no. 5, pp. 813–827, 2001.

[11] K. Worden, G. Manson, and D. Allman, "Experimental validation of a structural health monitoring methodology: part I. Novelty detection on a laboratory structure," *Journal of Sound and Vibration*, vol. 259, no. 2, pp. 323–343, 2003.

[12] Y.-L. Zhou, E. Figueiredo, N. Maia, R. Sampaio, and R. Perera, "Damage detection in structures using a transmissibility-based Mahalanobis distance," *Structural Control and Health Monitoring*, vol. 22, no. 10, pp. 1209–1222, 2015.

[13] K. Worden, G. Manson, and N. R. J. Fieller, "Damage detection using outlier analysis," *Journal of Sound and Vibration*, vol. 229, no. 3, pp. 647–667, 2000.

[14] A. Malhi and R. X. Gao, "PCA-based feature selection scheme for machine defect classification," *IEEE Transactions on Instrumentation and Measurement*, vol. 53, no. 6, pp. 1517–1525, 2004.

[15] A. Santos, E. Figueiredo, M. Silva, C. Sales, and J. Costa, "Machine learning algorithms for damage detection: Kernel-based approaches," *Journal of Sound and Vibration*, vol. 363, pp. 584–599, 2016.

[16] N. M. M. Maia, A. P. V. Urgueira, and R. A. B. Almeida, "Whys and wherefores of transmissibility," in *Vibration Analysis and Control - New Trends and Developments*, D. F. Beltran-Carbajal, Ed. InTech, 2011, pp. 197–216.

[17] B. Schölkopf, A. Smola, and K.-R. Muller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[18] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*. ACM, 1992, pp. 144–152.

[19] J. Mercer, "Functions of positive and negative type, and their connection with the theory of integral equations," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 209, no. 441–458, pp. 415–446, 1909.

[20] S. S. Keerthi and C.-J. Lin, "Asymptotic Behaviors of Support Vector Machines with Gaussian Kernel," *Neural Computation*, vol. 15, no. 7, pp. 1667–1689, 2003.

[21] D. Widjaja, C. Varon, A. Dorado, J. A. K. Suykens, and S. V. Huffel, "Application of kernel principal component analysis for single-lead-ECG-derived respiration," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, pp. 1169–1176, 2012.

[22] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. New York, United States: Springer-Verlag, 2002.