

# Blind Source Separation Methods for Video-based Structural Dynamics: A Comparative Study

Thaisse Dias Paes, Moisés Felipe Silva, João C. W. A. Costa

**Abstract**—Operational modal analysis has been effectively used for video-based structural dynamics identification in recent years. With several different approaches, the ones based on the blind source separation strategy have received increased attention for identifying structural dynamics characteristics. This work presents a comparison study on the blind source separation techniques for videos of vibrating structures. Two dimensionality reduction techniques are used for video compression along with six source separation algorithms, resulting in twelve different frameworks tested over a laboratory cantilever beam structure. For specific algorithms, the results indicate that dimensionality reduction techniques play a major role in the mode estimation performance.

**Keywords**—Video-based Structural Dynamics, Blind Source Separation, Dimensionality Reduction, Modal Analysis.

## I. INTRODUCTION

Modal analysis plays an important role as a tool for understanding the vibration characteristics of structures. Traditional methods require some instrumentation, such as physically-attached wired or wireless sensors, to perform experimental or operational modal analysis [1]. That requirement, however, leads to mass-loading on lightweight structures. Additionally, they are costly and time-consuming to install and maintain due to the significant maintenance for cabling (wired sensors) or the periodic replacement of the energy supply (wireless sensors). For these reasons, non-contact measurements, mainly the video-based ones, have received intensive research efforts in the last few years. They are relatively low-cost, agile and provide a very high spatial resolution measurements (every pixel effectively becomes a measurement point). In this way, video-based structural dynamics identification methods have been successfully applied for modal analysis, especially combined with vision-based algorithms [2] [3] [4] [5] [6].

In particular, blind source separation (BSS) techniques are prominent in the field of structural dynamics, allowing to blindly perform output-only modal identification [7]. However, the application of dimensionality reduction algorithms over the number of measurements is commonly required when applying BSS algorithms. Amid these compression techniques, principal component analysis (PCA) and nonnegative matrix factorization (NNMF) stand out [8] [9]. As PCA is often applied to reduce the dimensionality of the data, its principal components are closely related to the modal components of the structure,

Thaisse Paes and João Costa are with the Institute of Technology, Federal University of Pará, Belém, e-mails: thaisse.paes@itec.ufpa.br, jweyl@ufpa.br Moisés Felipe Silva is with the Institute of Communication, Information and Perception Technologies, Santa'Anna School of Advanced Studies, Pisa, Italy, e-mail: m.mellodasilva@santannapisa.it

with the advantage of preserving the high spatial resolution vibration information [6]. NNMF also achieve dimensionality reduction and motion decomposition by blindly identifying independent components. For this, the sources must be statistically dependent under conditions that impose additional constrains as non-negativity and sparsity [10]. Later, after choosing one particular technique for dimensionality reduction, BSS is carried out for modal analysis.

Due to its popularity, several works have explored BSS methods for structural dynamics. Musafere (2015) explores damage detection using BSS integrated with time varying auto-regressive modeling. The second-order blind identification algorithm (SOBI) solves the blind source separation problem. The technique is employed first to obtain the mono-harmonic responses from the vibration data, and then the discrete changes in the natural frequencies identify the instant and severity of damage [11]. Cheng (2016) analyses two different BSS algorithms, independent component analysis (ICA) and SOBI, for the condition monitoring of dams. The SOBI-based modal identification is suited to determine the system orders, while ICA calculates modal features to detect structural damage and determine the location of it [12].

For scenarios using vision-based approaches, in particular, those whose techniques require no surface preparation, BSS also plays an important role. Yang (2017) and Martinez (2020) proposed an efficient novel operational modal analysis method combining multi-scale pyramid decomposition and the complexity pursuit (CP) algorithm as a BSS technique. The method manipulates the spatio-temporal pixels phases that encode the local structural vibration in the video measurement, being capable to blindly identifying the modal frequencies, damping ratios and high-resolution mode shapes [4] [2].

Silva (2020) approaches full field, high-resolution modal identification from video associating NNMF and BSS solved by CP. The method also explores the timing information present in the pixel time-series to infer spatial relationships and, thus, blindly estimate modal parameters. The main difference between the aforementioned approach and the ones using phase-based optical-flow is that the previous is directly applied over the raw pixel time-series without any pre-processing step, such as the optical flow-based phase estimation. By using nonnegative matrix factorization, a spatio-temporal decomposition is achieved over the pixel time-series directly [6]. Later, Silva (2021) proposed adapted its technique for applications involving the estimation of travelling waves and complex mode shapes from video measurements. That generalized technique improves the decomposition results of the previous techniques by using the Hilbert Transform for augmenting the video data

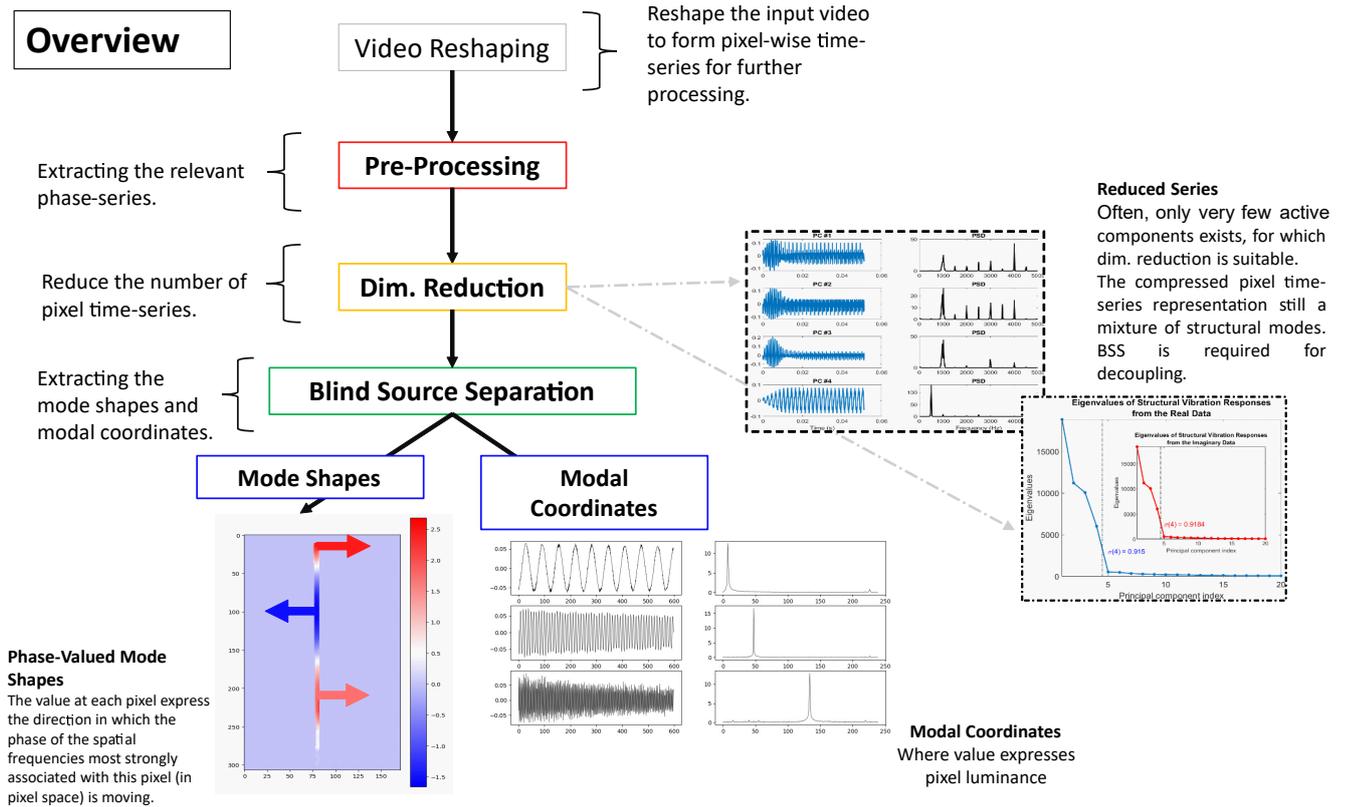


Fig. 1. Overview of the main steps involved in the framework for video-based structural dynamics estimation.

to obtain phase information. From the augmented dataset, PCA and BSS are applied to estimate real and imaginary modes, being capable of suitably decompose individual travelling wave sources into their fundamental components [13].

Although the literature addresses several BSS techniques, the CP algorithm is often employed for video-based modal analysis. The structural dynamics community, however, applies a variety of other BSS solutions that may be useful in this scope. This issue is tested here to provide a range of alternatives for the vide-based structural dynamics evaluation. This study explores and compares a set of BSS algorithms combined with dimensionality reduction methods for full-field high-resolution structural dynamics from video. The next section shows an outline of the method conducted for modal identification from video of vibrating structures and introduces the evaluated BSS techniques. Thereafter, the results are shown based on the extracted mode shape and modal coordinates.

## II. MODAL IDENTIFICATION FROM VIDEOS OF VIBRATING STRUCTURES

An overview of the methods applied for video-based structural dynamics identification is presented in this section. Figure 1 summarizes the output-only modal analysis framework for extracting full-field high-resolution modes from video of vibrating structures. The major three steps are pre-processing, dimensionality reduction and blind source separation.

The pre-processing component concerns the extraction of phase-valued time series for every pixel. Each time series maps

the temporal displacement for every pixel across the time-span of the video. Thereby, each pixel effectively becomes a measurement point and the number of time series is equals the resolution of the video. Later, the dimensionality reduction and blind source separation steps perform the structural system identification. This work investigates a subset of BSS techniques for modal analysis, reducing the number of pixel time series and identifying the mode shapes and modal coordinates.

### A. Dimensionality Reduction

Two techniques are applied interchangeably: PCA and NMF. Both are applied over all of the pixel time-series to perform a transformation from a high-dimensional space into a low-dimensional space. The low-dimensional space must retain the meaningful properties of the original data.

PCA reduces the number of input variables in a dataset into a small number of meaningful variables, called principal components. It is an amenable dimensionality reduction technique for orthogonal sources, especially when the most interesting sources have the largest variances. Even though it works great for reducing dimensionality, PCA *per se* is theoretically and empirically suboptimal for source separation [8].

In order to perform dimensionality reduction of the data from video of vibrating structures, Principal Component Analysis is applied across all of the time series at each pixel (not frames). It is expected the principal components to be closely related to the modal components, although not exactly

matching them. Ideally, the number of non-trivial eigenvalues calculated from the displacement time series is of the same order as the number of observable mode shapes. Therefore, the number of eigenvectors that correspond roughly to the number of nontrivial singular values is enough to retain the significant properties. Additionally, in this work, the PCA entry data was augmented by using Hilbert Transform to obtain both real and imaginary principal components.

Similarly, NNMF is another effective and attractive spatio-temporal method to reduce data dimension. NNMF decomposes a non-negative matrix  $\mathbf{V}$  into two non-negative matrices  $\mathbf{W}$  and  $\mathbf{H}$  whose product provides a good approximation to the original matrix. Due to its non-negative constraint, NNMF is extremely suitable for video processing (i.e., the pixel values are RGB intensities).  $\mathbf{W}$  acts as a dictionary of recurring patterns, regarded as the basis vectors. The dimensionality reduction is performed by projecting  $\mathbf{H}$  onto the lower dimensional space formed by basis vectors [9] [14].

For the dimensionality reduction of the data from video of vibrating structures problem, the entries of the NNMF algorithm are all the raw time-series from each pixel. It is assumed that the non-negative factors in  $\mathbf{W}$ , provided by the temporal decomposition process, are related to the modal components by indicating the time pattern of a specific vibration mode. The matrix  $\mathbf{H}$  incorporates a spatial decomposition and, when properly reshaped, resembles (but not exactly) the corresponding vibration mode shapes.

### B. Blind Source Separation

Blind source separation addresses the problem of recovering a set of source signals after they are linearly mixed by an unknown mixing matrix. The measured sensor signals are assumed to follow a linear superposition model. The problem is to find a separation matrix without knowing the mixing matrix and the source signals [15].

Using the dimension reduced components from PCA or NNMF, the BSS assumes, in the case of video-based structural dynamics applications, the structural vibration measured into the recorded video frames as a local time-varying motion from a temporally translated image intensity  $I(y + \delta(y, t))$ , where  $y$  is the pixel coordinate. The vibration motion is expressed in terms of a mixing mechanism, where  $S$  is a set of individual source signals (individual motions),  $s_i$  over time ( $t$ ), weighted by some mixing component  $A = [a_i]$  as

$$\delta(y, t) = \sum_{i=1}^k s_i(t) a_i = \mathbf{S}\mathbf{A}, \quad (1)$$

where,  $k$  is the number of vibration modes [6]. This is also expressed as a modal superposition, that is a linear combination of modal responses of an arbitrary structural system

$$\delta(y, t) = \mathbf{q}(t)\Phi(y) = \sum_{i=1}^k q_i(t)\phi_i(y), \quad (2)$$

where  $\mathbf{q}$  is the modal response matrix with  $q_i(t)$  being the  $i$ -th modal coordinate, and  $\Phi$  the mode shape matrix with each  $\phi_i(y)$  as the  $i$ -th mode shape. In the end, it should be noted

that the resulting BSS decomposition yields, directly, both the mode shapes and modal coordinates, with the later being the same as the sources estimated from the BSS. Other modal parameters, such as natural frequencies and damping ratios can be further estimated using conventional Fourier transform and logarithmic decrement techniques.

Several algorithms have been proposed in the literature to solve the BSS problem. Surrounded by a great number of BSS solutions, this work investigate six methods for video-based structural dynamics, namely: Complexity Pursuit (CP), Independent Component Analysis (ICA), Second Order Blind Identification (SOBI), SOBI Robust Orthogonalization (SOBIRO), Equivariant Robust ICA (ERICA), Algorithm for Multiple Unknown Signals Extraction (AMUSE) [15] [16][17].

### C. Methods for testing Dimensionality Reduction and BSS Algorithms

Twelve full-field high-resolution modal identification algorithms were proposed integrating PCA or NNMF to one of the six BSS techniques. The methods using principal component analysis are PCA-CP, PCA-ICA, PCA-SOBI, PCA-SOBIRO, PCA-ERICA and PCA-AMUSE. The methods in which the dimensionality reduction is based on nonnegative matrix factorization are NNMF-CP, NNMF-ICA, NNMF-SOBI, NNMF-SOBIRO, NNMF-ERICA and NNMF-AMUSE.

The modal identification algorithms were applied to an experimentally captured video of a cantilever. The structure was subjected to a single horizontal excitation caused by an impact hammer. A stationary camera was used for video recording at a frame rate of 480 frames per second. The video is 216×384 pixels (a total of 82944 pixels) with 600 frames captured at 480 frames per second. It is known from prior analysis that the cantilever beam exhibits 3 modes. The measured natural frequencies of these modes are 7.2 Hz, 47.2 Hz, and 133.6 Hz [12] [2].

## III. COMPARISON OF THE MODAL IDENTIFICATION RESULTS

A comparison of modal identification results is presented in this section. The twelve BSS-based methods attempt to identify the three vibration modes from the cantilever beam, along with their corresponding natural frequencies 7.2Hz, 47.2Hz and 133.6Hz. Figure 2 and Figure 3 give an outline of the PCA and NNMF-based approaches results, respectively. The figures show real and imaginary principal components with their power spectral density (PSD), the nonnegative factors, the de-mixing matrices for each method, the modal coordinates, the mode shapes, the modal assurance criteria (MACs) between the techniques and the damping ratios.

For the PCA-based approaches, PCA-CP, PCA-ICA, PCA-SOBIRO and PCA-ERICA identified all three modal frequencies. On the other hand, PCA-SOBI was not able to identify the third mode shape, with the natural frequency equivalent to 133Hz. PCA-AMUSE was not able to properly identify any of the three modes. In general, the first five methods seem to agree well on the natural frequency estimates of the three

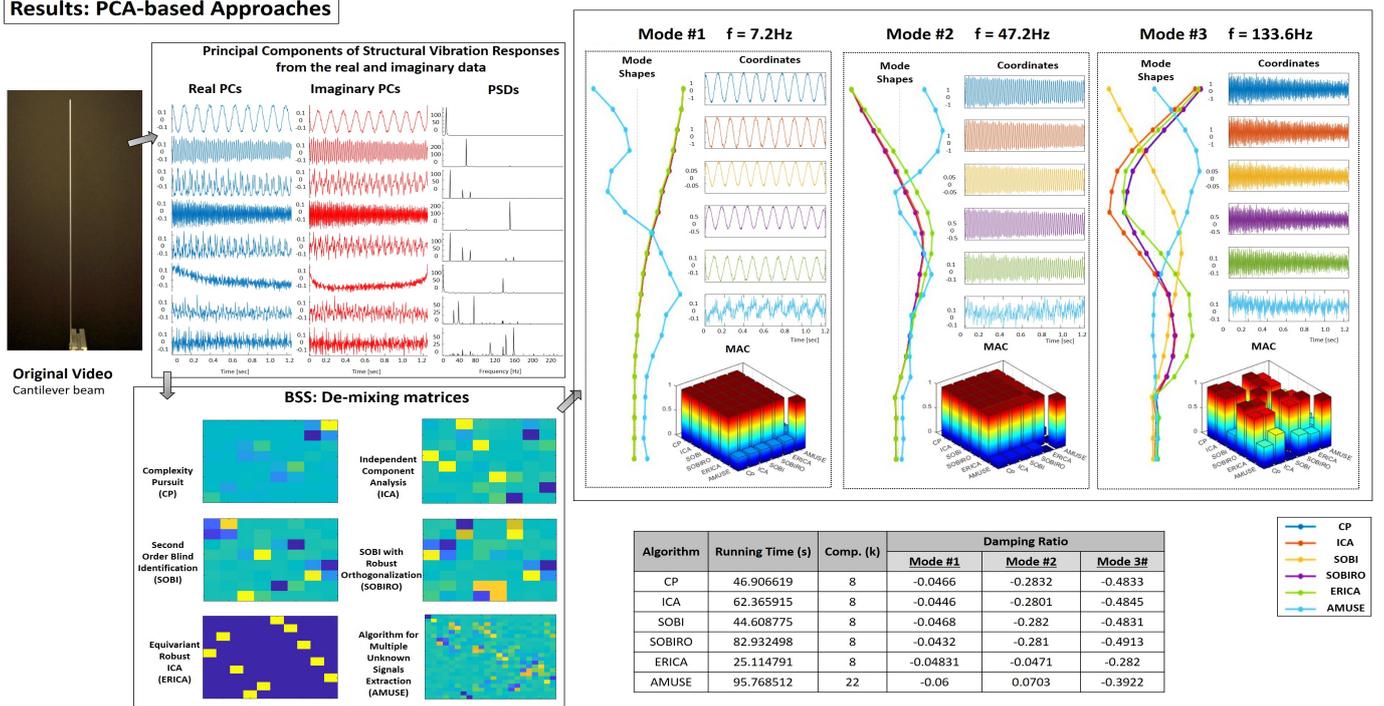
**Results: PCA-based Approaches**


Fig. 2. Demonstration of the PCA-based approaches results.

modes. The MAC's for each mode between the six techniques are shown graphically in Figure 2.

PCA-CP, PCA-ICA, PCA-SOBI, PCA-SOBIRO and PCA-ERICA presented monotone modal coordinates using the first eight principal components. PCA-AMUSE, on the other hand, used 22 principal components to present modal coordinates with monotone behavior (often, a properly estimated mode is presented as a monotone signal). Since the number of modes to be estimated from the video is three, and ideally, when using PCA along with the Hilbert transform for data augmentation, the minimum number of principal components is twice the number of modes possible to verify into the measurements. However, here, we observe that none of the techniques worked properly when using six principal components, instead, eight components were required for most techniques.

In summary, among the PCA-based methods, the most consistent for identifying the three modes are PCA-CP, PCA-ICA, PCA-SOBIRO and PCA-ERICA, although the third mode for the last method does not totally agree with the modes estimated using the other methods. PCA-SOBI and PCA-AMUSE are not indicated for the task of performing dynamics identification from video measurements, as they were not capable to estimate the three modes. More specifically, PCA-AMUSE could not estimate any of the modes, at any level; on the other hand, PCA-SOBI performed well on identifying the first two modes, but was not able to find the third mode.

For the NNMF-based methods, the NNMF-CP was the most successful in identifying the three mode frequencies properly. The first and second modes were identified by NNMF-CP, NNMF-SOBI and NNMF-SOBIRO. NNMF-ICA, NNMF-ERICA and NNMF-AMUSE were not able to properly identify the vibration modes, as can be seen by visualizing the

mode shapes and the MAC values shown in Figure 3.

NNMF-AMUSE was not successful in estimating any of the three modes, confirming its poor performance in performing modal analysis in the context of a video-based framework. For the remaining methods, NNMF-ICA and NNMF-ERICA also performed poorly when attempting to estimate any of the modes. Although PCA-SOBIRO was able to reasonably estimate the first two modes, it found problems for estimating the third one, mainly when considering the mode shapes. In general, when using NNMF as the dimensionality reduction technique, only CP and SOBI demonstrated to derive reasonable and consistent results in terms of mode shape estimation.

In general, for the present case, the BSS algorithms integrated into PCA presented better modal identification performance than the NNMF-based methods. It is found that CP properly performed modal analysis combined with both dimensionality reduction methods. ICA and ERICA only identified the modal shapes and mode coordinates combined with PCA. Couple with PCA, SOBI adequately identified the first two modes only, demonstrating a better performance when used along with the NNMF as it identified all modes. Regarding the SOBIRO-based method, it identified the three modes with PCA and only the first two modes combined with NNMF. AMUSE was not successful in accomplishing full-field high-resolution modal identification from video of vibrating structures in any case.

#### IV. CONCLUSIONS

As addressed, video-based methods successfully perform structural dynamics identification, with the advantage of requiring non-contact measurements. This work experimentally

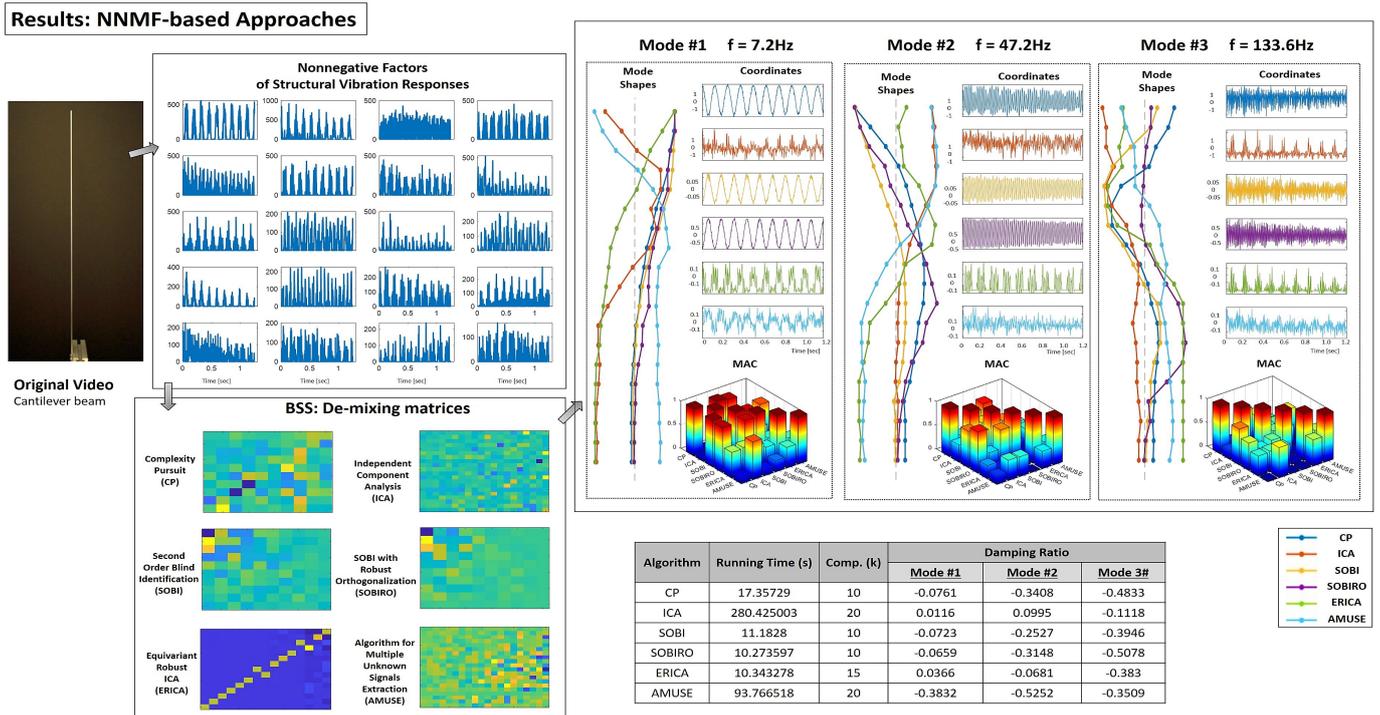


Fig. 3. Demonstration of the NNMF-based approaches results.

compared the performance of twelve different methods to extract three different modes from a cantilever beam video.

As these methods require BSS techniques, six different BSS solutions were investigated to analyze their applicability to perform modal estimation. From the results, the choice of the dimensionality reduction techniques plays an important role in the BSS algorithms performance, with only the CP-based methods estimating reasonable mode shapes when interchangeably used with the dimensionality reduction algorithms.

The results of this work contribute to the video-based structural dynamic field assisting the choice of the dimensionality reduction algorithms and providing a range of alternatives to the blind source separation-based approaches for video. As a future work, it is possible to investigate a certain method for specific cases of acquired images.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Professor Yongchao Yang, from the Michigan Technological University, and Dr David Mascareñas, from the Los Alamos National Laboratory, for providing us the video data. Also, the authors acknowledge the financial support received from CAPES and CNPq.

REFERENCES

[1] W. Fan, P. Qiao, Vibration-based damage identification methods: a review and comparative study, *Struc. Health Monit.*, v. 10, 83–111, 2011  
 [2] Yang, Y.; Dorn, C.; Mancini, T.; Talken, Z.; Kenyon, G.; Farrar, C.; Mascareñas, D. Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification. *Mechanical Systems and Signal Processing*. 85. 567-590. 2017  
 [3] J. Chen, N. Wadhwa, Y. Cha, F. Durand, W. Freeman, O. Buyukozturk, Modal identification of simple structures with high-speed video using motion magnification, *J. Sound Vib.* 345,58–71.2015

[4] Martinez B, Green A, Silva MF, Yang Y, Mascareñas D. Sparse and Random Sampling Techniques for High-Resolution, Full-Field, BSS-Based Structural Dynamics Identification from Video. *Sensors*, 2020  
 [5] Yang, Y.; Nagarajaiah, S.. Blind modal identification of output-only structures in time-domain based on complexity pursuit. *Journal of Earthquake Eng. & Struc. Dyn.*, 1885-1905, October, 2013  
 [6] Silva, M.; Martinez, B.; Figueiredo, E.; Costa, J.; Yang, Y.; Mascareñas, D.. Nonnegative matrix factorization-based blind source separation for full-field and high-resolution modal identification from video. *Journal of Sound and Vibration*, Volume 487, 2020  
 [7] Ganesh R. Naik and Wenwu Wang. Book: *Blind Source Separation: Advances in Theory, Algorithms and Applications (1st. ed.)*. Springer Publishing Company, Incorporated.2016.  
 [8] Jolliffe Ian T. and Cadima Jorge. Principal component analysis: a review and recent developments *Phil. Trans. R. Soc.* 374, 2016  
 [9] Y. Wang and Y. Zhang, "Nonnegative Matrix Factorization: A Comprehensive Review," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 6, pp. 1336-1353, June, 2013  
 [10] A. Cichocki, R. Zdunek, S. Amari, *New algorithms for non-negative matrix factorization in applications to blind source separation*, IEEE Inter. Conf. on Acoust. Speech and Sig. Proc., v. 5, 2006  
 [11] Musafere, F.; Sadhu, A.; Liu, K.. Towards damage detection using blind source separation integrated with time-varying auto-regressive modeling. *Smart Materials and Structures*, 015013, 2015  
 [12] Cheng, L.; Tong, F. Application of Blind Source Separation Algorithms and Ambient Vibration Testing to the Health Monitoring of Concrete Dams. *Math. Prob. in Eng.*, Hindawi, v. 2016, 1-15, December, 2016  
 [13] Silva, M.; Green, A.; Mascareñas, D.; Costa, J.; Figueiredo, E.. A Generalized Technique for Full-field Blind Identification of Travelling Waves and Complex Modes from Video Measurements with Hilbert Transform. *IMAC - Intern. Mod. Anal. Conf.*, Orlando, USA.2021  
 [14] Z. Yang, G. Zhou, S. Ding, S. Xie, Blind source separation by nonnegative matrix factorization with minimum-volume constraint *International Conf. on Intelligent Control and Inf. Processing*, pp. 117–119, 2010,  
 [15] Sadhu, A., Narasimhan, S., Antoni, J. A review of output-only structural mode identification literature employing blind source separation methods. *Mechanical Systems and Signal Processing*, 94, 415–431, 2017.  
 [16] Stone, James. Blind Source Separation Using Temporal Predictability. *Journal of Neural computation* vol. 13, 1559-74, August, 2001  
 [17] A. Delorme, J. Plamer, R. Oostenveld, J. Onton, and S. Makeig., "Comparing results of algorithms implementing blind source separation of EEG data.," *Swartz Foundation and NIH Grant*, 2007.