engineering proceedings

MDPI

Proceeding Paper

Full-Field Modal Analysis Using Video Measurements and a Blind Source Separation Methodology †

Samira Azizi [1,2,*], Kaveh Karami [1] and Stefano Mariani [2]

[1] Department of Civil Engineering, University of Kurdistan, Sanandaj 6617715175, Iran; ka.karami@uok.ac.ir
[2] Department of Civil and Environmental Engineering, Politecnico di Milano, 20133 Milano, Italy; stefano.mariani@polimi.it
* Correspondence: samira.azizi@polimi.it
† Presented at the 10th International Electronic Conference on Sensors and Applications (ECSA-10), 15–30 November 2023; Available online: https://ecsa-10.sciforum.net/.

Abstract: The adoption of wireless sensor networks has brought a significant breakthrough in structural health monitoring, providing an effective alternative to the challenges associated with traditional cable-based sensors. In recent years, a growing interest in developing contactless, vision-based vibration sensors like video cameras has led to advancements, potentially alleviating the previously mentioned drawbacks. In this study, a video of a vibrating frame is converted into a set of frames, so that local phase information can be extracted. The motion matrix is then derived from the phase information; since the number of measuring points is usually greater than the number of the excited modes of the system, the problem can become over-determined. Therefore, by applying dimensionality reduction techniques, the dimension of the motion matrix is significantly reduced. Finally, by exploiting an output-only identification technique, modal parameters are computed. The proposed approach is proven to accurately identify the structural frequencies and mode shapes.

Keywords: structural health monitoring; digital cameras; modal analysis; non-negative matrix factorization

Citation: Azizi, S.; Karami, K.; Mariani, S. Full-Field Modal Analysis Using Video Measurements and a Blind Source Separation Methodology. *Eng. Proc.* **2023**, *58*, 105. https://doi.org/10.3390/ecsa-10-16199

Academic Editor: Stefan Bosse

Published: 15 November 2023

1. Introduction

Modal analysis can be used to extract parameters that characterize the dynamic response of structures to external excitations [1]. This process can rely upon both input and output data or output data only. As the modal parameters depend on the inherent properties of the structure and are supposed to be unaffected by the external loads, they can be used for the assessment of the health of the monitored structures [2,3].

Typically, experimental and operational modal analyses involve the use of wired or wireless sensor networks, featuring optimized placement and/or cost [4–6], to observe the structural response [7,8]. Such sensors can provide measurements at sparse locations, often resulting in poor spatial resolution. In recent years, the capabilities of high-speed video measurements have significantly increased so that vision-based measurements can be effectively utilized in a wide range of applications in civil engineering, see, e.g., [9]. Estimating the optical flow by assuming a constant intensity between frames is a widely used technique; the equation imposing the invariant condition can be solved using different methods [10]. For instance, Chen et al. [11] identified the modes of cantilever beams and pipes by using phase-based motion magnification. Yang et al. [12,13] extracted the displacement of a three-story structure by combining phase-based optical flow and deep learning [14], while Dasari et al. [15,16] addressed the extraction of mode shapes and frequencies in the case of non-ideal rigid body motion. Martinez et al. [17] investigated the combination of compressed sensing sampling with system identification through images.

In the process of identification through video-based methods, due to the substantial dimensions of data, dimension reduction is necessary to reduce the dimensionality of the

motion matrix to the number of excited modes. This issue represents one of the main challenges within the field of vision-based identification.

In this study, we attack the problem of dimensionality reduction in two ways: first, full-field identification is achieved by handling a few pixels in the images; second, the dimensionality of said matrix is reduced to match the number of excited modes and, simultaneously, denoise the information. After having extracted the phase and amplitude from all of the frames of a video by applying the Gabor filter, pixels are selected by means of two criteria: a customarily adopted one, based on the amplitude, and a newly proposed one, based on the amplitude coupled with a Canny edge detection method. A motion matrix representing the time history of the pixel phase is then extracted and, via non-negative matrix factorization (NMF), the dimension of the data is reduced. Finally, the dimension of the phase matrix is reduced, and by means of complexity pursuit (CP), the full-field mode shapes and the vibration frequencies are obtained. The accuracy of the results is assessed with an example related to a three-story building.

2. Phase Extraction with the Gabor Filter

Phase information has been proven to be more correlated with structural motion than the amplitude and intensity value of the image [11]. In Fourier analysis, time delay in a signal corresponds to phase variations in the frequency domain; similarly, in two-dimensional signals (images), the spatial motion results in a phase change. In this context, we refer to the local phase corresponding to a pixel in a specific coordinate system, which can be extracted using the Gabor filter. This filter type, which is a complex quadrature filter, is the product of the Gaussian kernel and a complex sinusoidal function; due to article length constraints, additional details are omitted herein and readers are referred to [18]. The real and imaginary part of this filter in two directions is represented in Figure 1.

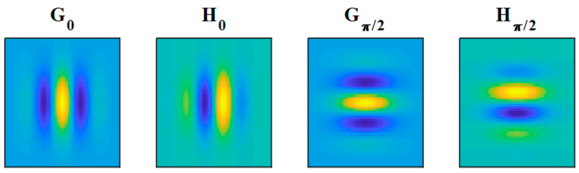


Figure 1. A pair of filters with a 90-degree phase difference, oriented horizontally and vertically in the image.

Assuming that $I(x, y, t_0)$ is the intensity of the image at time t_0 and at the location (x, y) in the frame, the local phase (Φ_θ) and local amplitude (A_θ) in the θ direction are obtained by convolving the filter $(G_\theta + iH_\theta)$ with the image as:

$$A_\theta(x, y, t_0)e^{i\Phi_\theta(x,y,t_0)} = (G_\theta + iH_\theta) \otimes I(x, y, t_0) \quad (1)$$

In the phase-based method, it is assumed that the local phase contour remains constant. The mathematical representation of this concept is given by:

$$\Phi_\theta(x, y, t_0) = C \quad (2)$$

This equation can be used to determine the displacement field.

3. Full-Field Identification

By applying the Gabor filter to the frames of a video of a vibrating structure, the local phase $\Phi(x, y, t)$ is obtained to capture the structural vibrations $d(x, y, t)$ provided as a linear combination of modal responses:

$$d(x, y, t) = \sum_{i=1}^{n} \varphi_i(x, y) q_i(t) \quad (3)$$

where n is the number of excited modes; $\phi \in R^{N\times n}$ is the matrix of the mode shapes, with φ_i being the i-th vibration mode; N is the number of pixels in one frame; and $q(t) \in R^{n\times T}$ is the vector of modal coordinates, with T being the number of handled frames.

Within blind source separation (BSS) methodologies aimed at modal identification, modal responses and vibration frequencies are extracted from the measured output without prior knowledge of the mixing matrix that represents the mode shapes. This is accomplished via the measured output by extracting phase information from the frames. Given that the number of image pixels is significantly higher than the number of typically excited modes, namely given that N >> n, the problem becomes characterized by a large amount of data leading to high computational costs. Consequently, the direct application of BSS methods is not feasible for modal identification; dimensionality reduction techniques must be employed to reduce the dimensionality of the time-series matrix.

3.1. Non-Negative Matrix Factorization

Non-negative matrix factorization was introduced in [15]. Assuming the matrix $V \in R^{n\times m}$ to comprise the phase time history of pixels (motion matrix), reduced to rank r, its non-negative factorization involves the two matrices $W \in R^{n\times r}$ and $H \in R^{r\times m}$ such that $V \approx WH$. This is achieved by minimizing the Euclidean distance between V and WH, according to:

$$\min \ \|V - WH\| \ \text{ subjected to } \ W \geq 0, H \geq 0 \tag{4}$$

NMF relies upon non-convex optimization during its iterative process, potentially resulting in different solutions based on the initialization values of W and H. The initialization procedure adopted in this paper is as follows: the eigenvectors corresponding to the largest eigenvalues (larger than 1% of the maximum one), as a result of PCA on the pseudo inverse of the motion matrix, are retained in the analysis; W is obtained by applying ICA on the whitened matrix of step 1; H is obtained by means of the motion matrix. Not only can different initialization procedures to set W and H affect the accuracy of the identified modes but also the reduced dimension of the problem can be modified; this is a rather common challenge in vision-based identification problems. A comparison between the adopted method and others available in the literature is beyond the scope of this conference paper.

3.2. Complexity Pursuit

By applying NMF on the extracted phase matrix, the reduced one W can be obtained. Assuming that W', where the prime stands for transpose, can be decoupled into modal coordinate q according to:

$$q = LW' \tag{5}$$

The de-mixing matrix $L^{r\times r}$ has to be acquired through the use of CP [19]. The mode shape matrix, see Equation (3), is then computed with:

$$\phi = H'L^{-1} \tag{6}$$

4. Numerical Example

In this study, we assess the capability of the procedure outlined in Section 3, by means of an on-purpose generated video of a three-story structure with mass $m = [12 \ \ 10 \ \ 5] \times 10^3$ kg, stiffness $k = [45 \ \ 40 \ \ 23] \times 10^6$ N/m, and damping proportional to the mass, subject to free vibrations. The phase matrix is obtained by applying the Gabor filter to all of the frames. To effectively handle the data volume and to prove the feasibility of obtaining mode shapes using a reduced number of pixels, the procedure reads as follows: pixels are selected on the basis of the amplitude in the initial frame; Canny edge detection is employed in conjunction with the amplitude information; and NMF and CP are applied to the phase matrix. The entire procedure is sketched out in Figure 2.

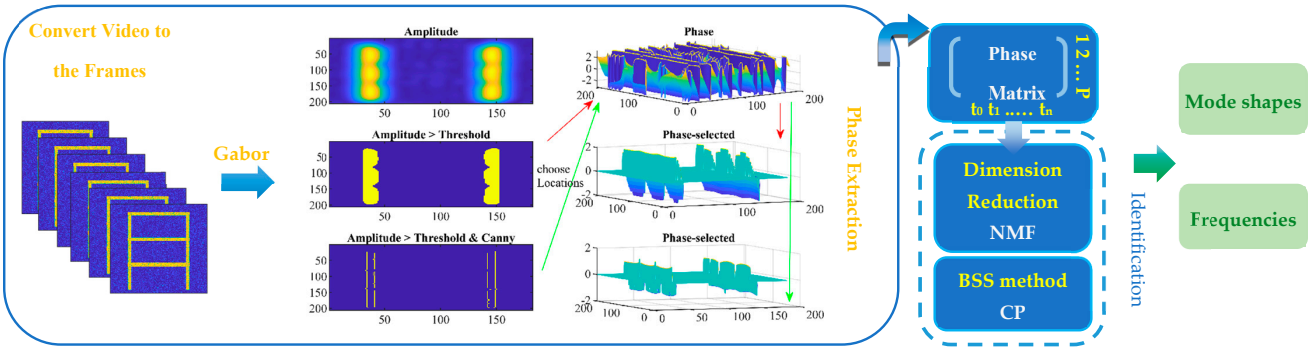
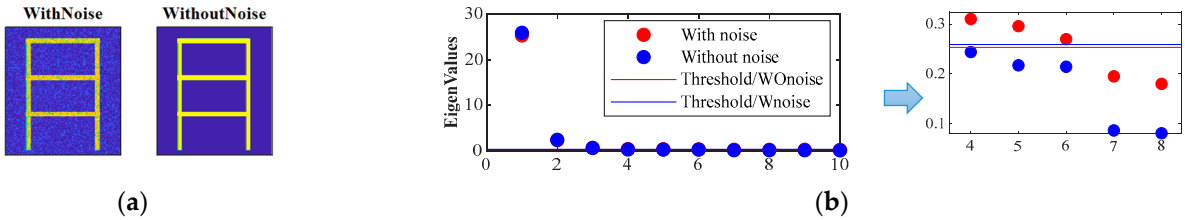


Figure 2. Schematic diagram of the proposed NMF−CP based procedure.

As previously mentioned, one of the inherent challenges within vision-based identification methodology pertains to dimension reduction. At this stage, obtaining a matrix with dimensions matching the number of excited modes of a vibrating structure is not straightforward. To illustrate this issue, we refer to the approach employed in [20] for a three-story structure, whose motion to catch the vibrations can be affected by noise, see Figure 3a. A singular value decomposition (SVD) of the covariance matrix derived from the phase matrix is adopted. The dimension reduction is based on the selection of the eigenvalues exceeding 1% of the largest one, as depicted in Figure 3b. If noise is present, the number of eigenvalues exceeding this threshold turns out to be three.



(**a**) (**b**)

Figure 3. (**a**) A frame of the video with and without noise; (**b**) eigenvalues of the covariance of the phase matrix.

Upon the adoption of CP for the phase matrix, the number of identified modes may be larger than the actual number of excited modes, thereby needing a post-processing step. While it is relatively straightforward to visually select the correct modes in structures with a simple geometry, it may become prohibitively difficult for structures characterized by a complex shape.

By applying the procedure illustrated in Figure 2 to the noisy video, the phase matrix is obtained using two distinct phase selection methods. The dimension of the phase matrix obtained by employing the Canny edge detection in combination with the amplitude information results in it being smaller than the one obtained by solely using the amplitude data. Figure 4 shows that the identified mode shapes are identical in the two cases.

In Figure 5, a comparison between the identified mode shapes extracted from the video and the mode shapes used to generate the video is provided. It is evident that the first and second modes closely align with the real mode shapes of the structure at story height. However, the third mode shows some disparities at the first story. Upon comparison with the actual values obtained using the modal assurance criterion, it can be observed that the identified vibration frequencies exhibit a high degree of similarity with the real ones, amounting to at least 98.2%.

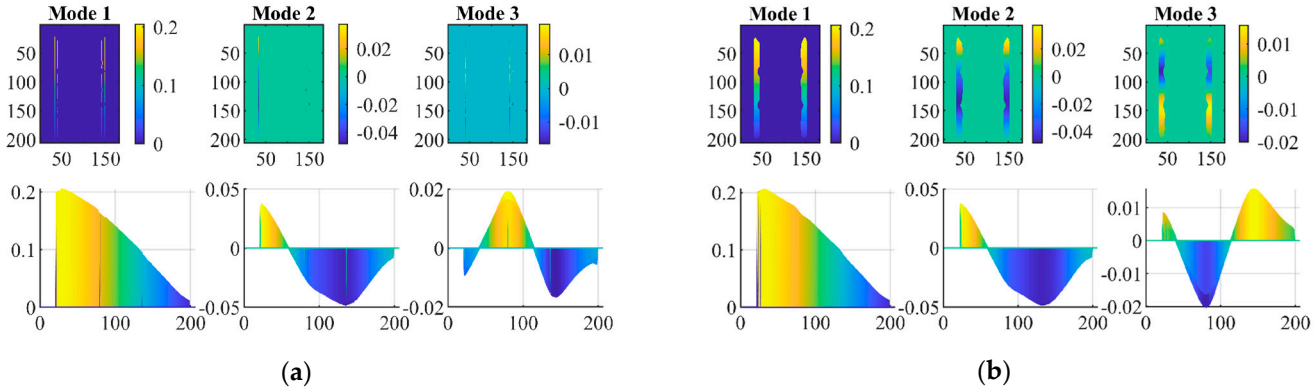


Figure 4. Identified mode shapes using: (**a**) the Canny and amplitude technique and (**b**) the amplitude technique only.

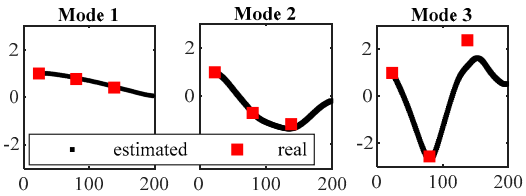


Figure 5. Identified mode shapes.

5. Conclusions

This study has focused on the identification of the mode shapes and frequencies of vibrating structures. Given the substantial dimensionality of the data to handle such aspects in the analysis, we implemented a two-step pixel selection procedure by applying a threshold to the pixel amplitude obtained by way of the Canny edge detection method. Next, we applied NMF to reduce the dimensionality of the phase matrix and to also denoise the phase time-series. In fact, it has been demonstrated that noisy data can significantly impact the eigenvalue selection phase, potentially leading to the identification of incorrect mode shapes. Subsequently, CP was employed for the identification process.

The results show that this approach can be an accurate and reliable method to identify vibration frequencies and mode shapes, at least for problems related to the vibrations of shear-type buildings. The denoising procedure in fact resulted in the accurate identification of mode shapes. In future research, we aim to investigate the impact of the initialization method of NMF and to refine the process of selecting the rank for NMF dimension reduction.

Author Contributions: Conceptualization, S.A., S.M. and K.K.; methodology, S.A. and K.K.; validation, S.A., S.M. and K.K.; formal analysis, S.A.; resources, K.K.; data curation, S.A.; writing—original draft preparation, S.A.; writing—review and editing, S.M.; visualization, S.M.; supervision, S.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fu, Z.-F.; He, J. *Modal Analysis*; Elsevier: Amsterdam, The Netherlands, 2001.
2. Torzoni, M.; Rosafalco, L.; Manzoni, A.; Mariani, S.; Corigliano, A. SHM under varying environmental conditions: An approach based on model order reduction and deep learning. *Comput. Struct.* **2022**, *266*, 106790. [CrossRef]
3. Azizi, S.; Karami, K.; Nagarajaiah, S. Developing a semi-active adjustable stiffness device using integrated damage tracking and adaptive stiffness mechanism. *Eng. Struct.* **2021**, *238*, 112036. [CrossRef]
4. Capellari, G.; Chatzi, E.; Mariani, S.; Azam, S.E. Optimal design of sensor networks for damage detection. *Procedia Eng.* **2017**, *199*, 1864–1869. [CrossRef]
5. Bruggi, M.; Mariani, S. Optimization of sensor placement to detect damage in flexible plates. *Eng. Optim.* **2013**, *45*, 659–676. [CrossRef]
6. Capellari, G.; Chatzi, E.; Mariani, S. Cost-benefit optimization of sensor networks for SHM applications. *Proceedings* **2018**, *2*, 132.
7. Liu, Y.; Bao, Y. Real-time remote measurement of distance using ultra-wideband (UWB) sensors. *Autom. Constr.* **2023**, *150*, 104849. [CrossRef]
8. Liu, Y.; Liu, L.; Yang, L.; Hao, L.; Bao, Y. Measuring distance using ultra-wideband radio technology enhanced by extreme gradient boosting decision tree (XGBoost). *Autom. Constr.* **2021**, *126*, 103678. [CrossRef]
9. Spencer, B.F., Jr.; Hoskere, V.; Narazaki, Y. Advances in computer vision-based civil infrastructure inspection and monitoring. *Engineering* **2019**, *5*, 199–222. [CrossRef]
10. Chen, J.G.; Adams, T.; Sun, H.; Bel, E.S.; Büyüköztürk, O. Camera-based vibration measurement of the world war I memorial bridge in Portsmouth, New Hampshire. *J. Struct. Eng.* **2018**, *144*, 04018207. [CrossRef]
11. Chen, J.G.; Davis, A.; Wadhwa, N.; Durand, F.; Freeman, W.; Büyüköztürk, O. Video camera–based vibration measurement for civil infrastructure applications. *J. Infrastruct. Syst.* **2017**, *23*, B4016013. [CrossRef]
12. Fleet, D.J. *Measurement of Image Velocity*; Springer Science & Business Media: Berlin, Germany, 2012; Volume 169.
13. Chen, J.G.; Davis, A.; Wadhwa, N.; Cha, Y.-J.; Durand, F.; Freeman, W.T.; Buyukozturk, O. Modal identification of simple structures with high-speed video using motion magnification. *J. Sound Vib.* **2015**, *345*, 58–71. [CrossRef]
14. Luan, L.; Zheng, J.; Wang, M.L.; Yang, Y.; Rizzo, P.; Sun, H. Extracting full-field subpixel structural displacements from videos via deep learning. *J. Sound Vib.* **2021**, *505*, 116142. [CrossRef]
15. Silva, M.; Martinez, B.; Figueiredo, E.; Costa, J.C.; Yang, Y.; Mascareñas, D. Nonnegative matrix factorization-based blind source separation for full-field and high-resolution modal identification from video. *J. Sound Vib.* **2020**, *487*, 115586. [CrossRef]
16. Dasari, S.; Dorn, C.; Yang, Y.; Larson, A.; Mascarenas, D. A framework for the identification of full-field structural dynamics using sequences of images in the presence of non-ideal operating conditions. *J. Intell. Mater. Syst. Struct.* **2018**, *29*, 3456–3481. [CrossRef]
17. Martinez, B.; Green, A.; Silva, M.F.; Yang, Y.; Mascareñas, D. Sparse and random sampling techniques for high-resolution, full-field, bss-based structural dynamics identification from video. *Sensors* **2020**, *20*, 3526. [CrossRef] [PubMed]
18. Freeman, W.T.; Adelson, E.H. The design and use of steerable filters. *IEEE Trans. Pattern Anal. Mach. Intell.* **1991**, *13*, 891–906. [CrossRef]
19. Yang, Y.; Nagarajaiah, S. Blind modal identification of output-only structures in time-domain based on complexity pursuit. *Earthq. Eng. Struct. Dyn.* **2013**, *42*, 1885–1905. [CrossRef]
20. 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. *Mech. Syst. Signal Process.* **2017**, *85*, 567–590. [CrossRef]