Vision-based operational modal analysis robust to environmental conditions

Zhilei Luo, Boualem Merainani, Michael Döhler, Vincent Baltazart, Qinghua Zhang

Vision-based operational modal analysis robust to environmental conditions

Zhilei Luo[1], Boualem Merainani[2], Michael Döhler[1], Vincent Baltazart[2], and Qinghua Zhang[1]

[1] Université Gustave Eiffel, Inria, COSYS-SII, I4S Team, F-35042 Rennes, France
[2] Université Gustave Eiffel, Inria, COSYS-SII, I4S Team, F-44344 Bouguenais, France

Abstract. Vision measurements are becoming a powerful technique for operational modal analysis. Nevertheless, the performance of the existing motion extraction algorithms is subject to perturbations due to environmental interference in practical applications, such as illumination variations and objects in the background of the investigated structure. This paper applies a novel image-based motion estimation method that explicitly considers such perturbations, and compares the performance of different motion extraction methods with regards to the modal parameter estimates. These methods are tested on video data recorded from a cantilever beam in the laboratory under random base excitation, where the robustness against ambient light changes and disturbance by background features is investigated. The goal of this work is to move towards real-world applications of vision-based operational modal analysis by improving robustness.

Keywords: Vision measurement · motion extraction · light changes · subspace methods · operational modal analysis.

1 Introduction

Vision-based measurements have developed into a powerful full-field inspection tool for engineering structures. Over the last decade, such non-contact measurement methods have succeeded in various applications of vibration-based operational modal analysis (OMA), especially for civil and mechanical structure monitoring [5, 3, 4].

With vision measurements, the vibrating behavior of an object is captured and recorded into video frames, characterized by image intensity features. Vision-based motion estimation algorithms aim of comparing pixel intensity changes in video frames to extract displacements or velocities of moving object. Thus, for long span structures, it can easily provide hundreds or thousands of virtual sensors. Currently prevalent algorithms for vibration estimation include digital image correlation [16, 1], Kanade-Lucas-Tomasi (KLT) optical flow tracking [15, 18], or global and local phase correlation [11, 7]. After motion estimation, the obtained structural vibration responses can be processed by classical system identification methods for operational modal analysis (OMA), e.g. stochastic subspace

identification [8] or PolyMAX [14], to obtain the desired modal parameters from the measurements.

Since the recorded video data is processed to obtain vibration data for OMA, the precision of the modal parameter estimates hinges on the quality of the motion estimation algorithm [7]. Moreover, in real-world applications the illumination conditions may vary, i.e., brightness and contrast may change during a measurement. Hence, vision-based modal analysis is susceptible to such ambient condition changes if the motion estimation algorithm is not robust towards such changes [17].

This paper investigates the robustness of vision-based OMA to illumination variations, and a robust motion estimation method is proposed. The approach is compared with conventional techniques and validated on a laboratory specimen in a complex operational environment. Beyond considering changes in the illumination conditions, corrupting background features are taken into account by an active pixel selection for robust motion estimation.

2 Robust vision-based motion estimation

A robust two-step digital image correlation method is proposed for vibration measurement in this section, where illumination variations within a dataset are explicitly considered. Then, the computed displacements are used for modal analysis, e.g., using subspace identification [9, 2], where an efficient version for high-dimensional measurements [6] can be used.

2.1 Two-step image correlation

The digital image can be regarded as a 2D discrete signal $I(i, j, t)$, which is described by the spatial coordinates $i, j \in \mathbb{Z}$ and the temporal index of video frame t. First, regions-of-interest (ROIs) are selected in the image along the edges of the investigated structure, with the purpose of estimating the structural motion for each of the ROIs. The ROIs are small enough to assume no deformation of the investigated structure within the ROI, but only structural displacements. Within each ROI, the active pixels representing the target are automatically selected. The displacement of target occurs between frames. Denote the set of active pixels by $\mathbf{A}$, and then the pixel level displacement of the target is estimated by maximizing the zero mean normalized cross correlation (ZNCC) between reference frame at t_1 and displaced frame at t_2

$$(\hat{p}, \hat{q}) = \arg\max_{p,q} \frac{\sum_{i,j\in\mathbf{A}} I_z(i, j, t_1) I_z(i + p, j + q, t_2)}{\sqrt{\sum_{i,j\in\mathbf{A}} I_z(i, j, t_1)^2 \sum_{i,j} I_z(i + p, j + q, t_2)^2}}, \tag{1}$$

where (p, q) is the bidirectional pixel level displacement and $(\hat{p}, \hat{q})$ is the optimal value. $I_z(i, j, t)$ represents $I(i, j, t)$ after removing the mean value.

To simplify notation, the displaced frame is shifted to compensate the pixel level displacement

$$I_s(i, j, t_2) = I(i + \hat{p}, j + \hat{q}, t_2), \tag{2}$$

and subpixel correlation between frames is further carried out. Usually, interpolation techniques are applied to account for the inter-pixel image intensity [13]. Here, linear plane interpolation is employed on the reference frame

$$I(i+x, j+y, t_1) = a_{i,j}x + b_{i,j}y + c_{i,j}, \tag{3}$$

where (x, y) is the subpixel displacement to be estimated, and $a_{i,j}$, $b_{i,j}$ and $c_{i,j}$ are pixelwise coefficients for the linear interpolation. In this sense, the subpixel displacement between the current and the reference frames can be determined by solving the classical least squares problem

$$(\hat{x}, \hat{y}) = \arg\min_{x,y} \sum_{i,j \in \mathbf{A}} \left[I_s(i, j, t_2) - (a_{i,j}x + b_{i,j}y + c_{i,j}) \right]^2. \tag{4}$$

2.2 Robustness to illumination variation

Now, the effect of illumination variation on vision-based displacement estimation is considered. In fact, changes of brightness and contrast change the image intensity by [10]

$$\alpha I(i, j, t) + \beta, \tag{5}$$

where α and β are scalars. Substituting Eq. (5) to Eq. (4) and defining instrumental variables

$$\tilde{x} = \alpha x, \ \tilde{y} = \alpha y, \tag{6}$$

alter the formula of the original least squares problem into

$$(\hat{\tilde{x}}, \hat{\tilde{y}}, \hat{\alpha}, \hat{\beta}) = \arg\min_{\tilde{x}, \tilde{y}, \alpha, \beta} \sum_{i,j \in \mathbf{A}} \left[I_s(i, j, t_2) - (a_{i,j}\tilde{x} + b_{i,j}\tilde{y} + c_{i,j}\alpha + \beta) \right]^2, \tag{7}$$

where possible changes in brightness and contrast are explicitly considered. Its solution is easy to obtain, and the estimated subpixel displacement can be determined

$$\hat{x} = \frac{\hat{\tilde{x}}}{\hat{\alpha}}, \ \hat{y} = \frac{\hat{\tilde{y}}}{\hat{\alpha}}, \tag{8}$$

where $\hat{x}$, $\hat{y}$, $\hat{\tilde{x}}$, $\hat{\tilde{y}}$ and $\hat{\alpha}$ are the estimates. In this way, the effect of illumination variation is eliminated in the subpixel displacement estimation by Eq. (8).

3 Applications

In this section, the proposed motion estimation algorithm is illustrated on a cantilever beam structure in the Structures and Integrated Instrumentation Laboratory at Université Gustave Eiffel. The performance of the algorithm is compared with other conventional vision-based motion estimation algorithms:

- **ST.** Extended steerable filtering based large motion estimation [7].
- **POC.** Global phase correlation [11].

– **KLT.** Kanade-Lucas-Tomasi optical flow estimation, implemented in MATLAB [12].
– **Our method.** Proposed method in this paper robust to image brightness and contrast variations.

It should be noted that ST and POC should show some robustness towards illumination variation since they are phase-based methods.

3.1 Experimental setup

A steel cantilever beam is placed on a shake table, as presented in Fig. 1, and excited by a random noise that is band limited between 1 Hz and 200 Hz. The direction of excitation is perpendicular to the beam body. A camera is fixed in front of the beam so that the vibration direction of the beam is also perpendicular to the optical axis of camera. Laser sensors are installed beside the beam to provide a simultaneous displacement measurement for comparison at three different positions as shown in the right part of Fig. 1. The sampling rate of the camera is 600 frames per second, which makes it possible to estimate the first 8 modes of the beam from video data. Data is recorded for 2 minutes in the test, so 72000 video frames are generated in the test. During recording, two cases are considered: 1) simple condition, i.e. there is no illumination variation and no background features outside the target (see Fig. 2(a)); 2) the illumination of the experiment is modulated between 100% and 50% intensity of lamps at a frequency of 1 Hz, and the background is corrupted by a net-shaped structure to provide potentially disturbing features in the video frames (see Fig. 2(b)).

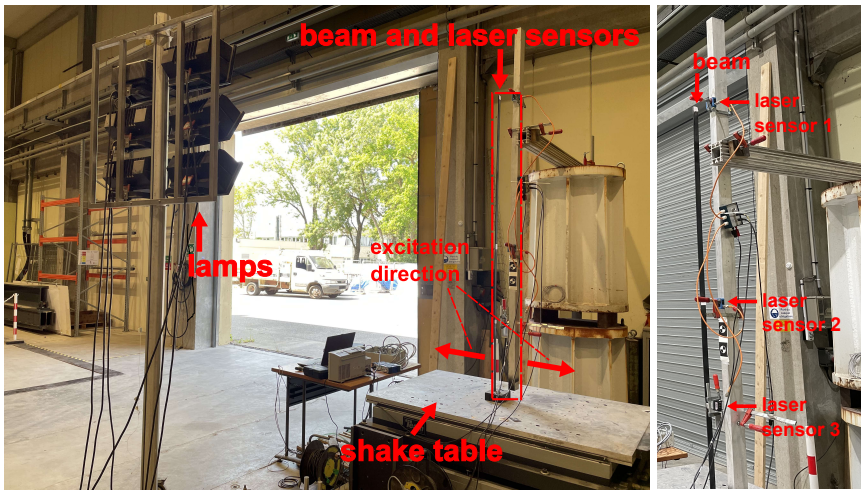

Fig. 1. Left: experimental setup. Right: Close-up on cantilever beam structure and laser sensors.

3.2 Motion estimation

The vibration estimation performance of vision based algorithms is assessed. In the video captured, the entire resolution of frame is 1632×128 pixels and 163 ROIs of size 10×20 pixels are set along the beam structure, as exhibited in Fig. 2(a) and (b). Among these ROIs, the displacements extracted from ROIs 1, 84 and 126 (marked red in Fig. 2) correspond to the laser measurement data. The differences are evaluated by the root mean square error (RMSE) in pixel unit:

$$E_{\mathrm{RMSE}} = \sqrt{\frac{\sum_{t=1}^{N_s}(\hat{u}_t - u_t)^2}{N_s}} \tag{9}$$

where t is the image frame index in the simulated image sequence. u_t and $\hat{u}_t$ are perpendicular displacements to beam body, which are measured by the laser and estimated by a motion extraction method respectively. N_s is the total number of image frames.

The RMSE results are presented in Table 1 for the three laser sensor positions. At these places, the error rate (in %) is calculated as a ratio between error level and maximum laser-measured displacement magnitude. The results show that the complex environmental conditions (i.e., illumination variations and objects in the background) indeed impair the vision-based motion estimation. At *Position 1*, the error level of ST and POC increases by more than a factor of 4. At *Position 2* and *Position 3*, the situation is similar with ST and POC but more moderate, possibly due to larger displacement and consequently a higher SNR at these positions. In contrast, our method is less impacted by the complex environmental conditions, with much less error increase. The KLT optical flow estimation fails to estimate the beam vibration in this test.

The computational time of algorithms is also recorded and exhibited in Table 1. Among tested algorithms, ST method generally costs the most time with

Table 1. Root mean square errors (RMSE) of beam vibration measurement, and computational time used for single frame in the video.

Evaluated methods	E_{RMSE}(pixel) and error rate (%)			Time used (milisecond/frame)
	Position 1	*Position 2*	*Position 3*	
simple condition				
ST	0.1004 (2.35%)	0.0572 (2.69%)	0.0249 (1.13%)	673.4
POC	0.1071 (2.50%)	0.0671 (3.15%)	0.0529 (2.42%)	71.0
KLT	0.3730 (8.72%)	No	No	36.0
Our method	0.1060 (2.48%)	0.0572 (2.69%)	0.0357 (1.63%)	61.0
illumination variation and unclean background				
ST	0.8817 (26.10%)	0.0575 (3.22%)	0.1329 (7.04%)	652.1
POC	0.4511 (13.35%)	0.1139 (6.37%)	0.1262 (6.69%)	72.6
KLT	No	No	0.2781 (14.74%)	34.4
Our method	0.2921 (8.65%)	0.0575 (3.22%)	0.0988 (5.24%)	67.8

more than 650 milliseconds per frame. POC is second in place with around 71 milliseconds per frame, to which our method is comparable with 61-68 milliseconds per frame. KLT optical flow estimation is the fastest, but fails to estimate the motions in our case.

3.3 Modal Analysis

The modal analysis accuracy of vision measurement method ST, POC and our method is evaluated. The KLT method is excluded due to its failure in vibration

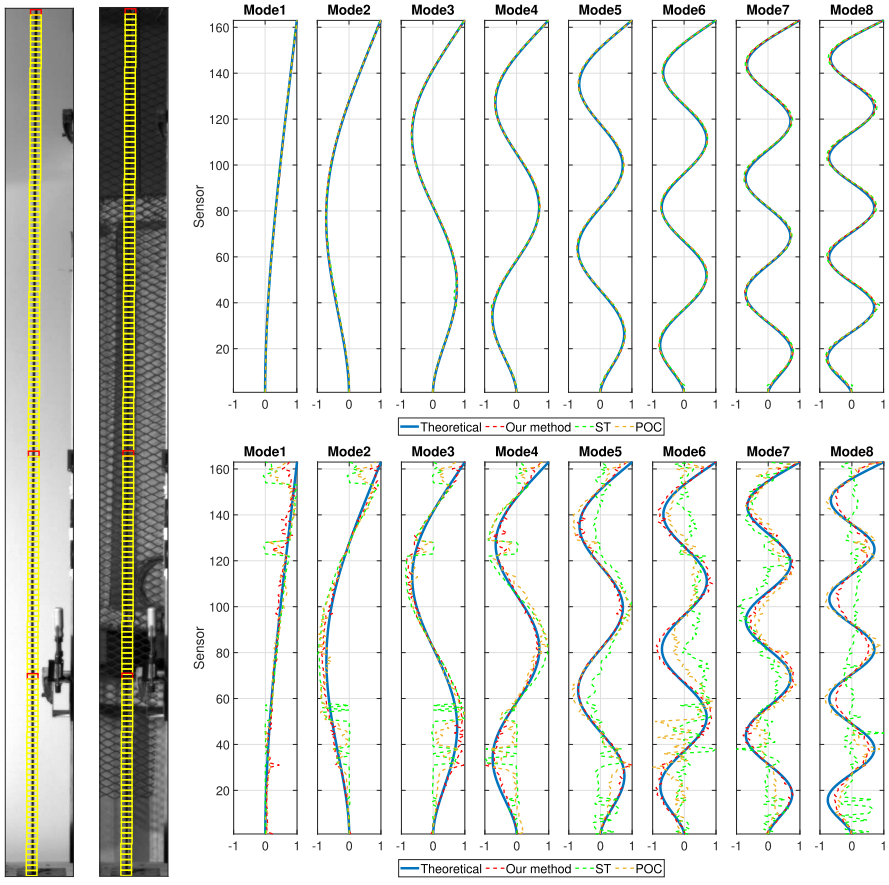

Fig. 2. Video frame including 163 regions of interest (ROI) with (a) clean and (b) unclean background. The red boxes are ROIs corresponding to laser sensing positions, while others without laser are marked yellow. (c) First 8 mode shapes of beam obtained by finite element analysis and vision algorithms in conditions without (upper) and with (lower) illumination variation and unclean background.

Table 2. Modal assurance criterion (MAC) values between mode shapes estimated from vision measurements and mode shapes from numerical model.

Algorithms	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6	Mode 7	Mode 8
simple condition								
ST	0.9999	0.9990	0.9989	0.9989	0.9982	0.9976	0.9963	0.9950
POC	0.9997	0.9992	0.9997	0.9992	0.9986	0.9985	0.9978	0.9968
Our method	0.9999	0.9995	0.9996	0.9993	0.9988	0.9986	0.9978	0.9969
illumination variation and unclean background								
ST	0.7316	0.7408	0.5489	0.5514	0.3625	0.0506	0.4107	0.0185
POC	0.8853	0.9013	0.8465	0.6976	0.8908	0.3749	0.7111	0.8864
Our method	0.9632	0.9906	0.9835	0.9787	0.9834	0.9862	0.9692	0.9784

estimation. The modal parameters are estimated using subspace identification [9, 2, 6]. The influence of the employed motion estimation method on the frequencies and damping ratios seems to be minor, so we focus on the mode shapes in the evaluation for sake of brevity. They are compared with theoretical mode shapes by finite element analysis in Fig. 2 and in Table 2, showing the MAC values.

In the simple case (no illumination variation and clean background), the mode shapes are very well estimated with either of the considered motion estimation methods, having MAC values all above 0.995. However, for the complex case with illumination variation and unclean background, apparent distinctions appear between the results from our method and ST and POC. Compared with the simple condition, the mode shapes based on ST and POC are significantly more corrupted than with our method in Fig. 2(c) (bottom). This is also quantitatively indicated by the lower MAC values with ST and POC, while the MAC values of our method are all higher than 0.96.

4 Concluding remarks

This paper demonstrates the importance of robust motion estimation for vision-based modal analysis. It has been illustrated that classical motion estimation methods like ST and POC may perform poorly under realistic application conditions where illumination variations may appear and other objects than the investigated structure are present in the background of the recorded video. This motivates the explicit consideration of such conditions in the motion estimation method, prior to the modal analysis. An efficient interpolation-based method with these properties has been introduced in this paper, and its better performance under such complex environmental conditions has been illustrated in a lab experiment.

References

1. Cai, E., Zhang, Y., Lu, X., Li, P., Zhao, T., Lin, G., Guo, W.: A target-free video structural motion estimation method based on multi-path optimization. Mechani-

cal Systems and Signal Processing **198**, 110452 (2023)
2. Döhler, M., Mevel, L.: Fast multi-order computation of system matrices in subspace-based system identification. Control Engineering Practice **20**(9), 882–894 (2012)
3. Feng, D., Feng, M.Q.: Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection – A review. Engineering Structures **156**, 105–117 (2018)
4. Fu, Y., Shang, Y., Hu, W., Li, B., Yu, Q.: Non-contact optical dynamic measurements at different ranges: a review. Acta Mechanica Sinica **37**(4), 537–553 (2021)
5. Janeliukstis, R., Chen, X.: Review of digital image correlation application to large-scale composite structure testing. Composite Structures **271**, 114143 (2021)
6. Luo, Z., Merainani, B., Döhler, M., Baltazart, V., Zhang, Q.: High dimensional data reduction in modal analysis with stochastic subspace identification. In: Proc. 22nd IFAC World Congress. Yokohama, Japan (2023)
7. Merainani, B., Xiong, B., Baltazart, V., Döhler, M., Dumoulin, J., Zhang, Q.: Subspace-based modal identification and uncertainty quantification from video image flows. Journal of Sound and Vibration **569**, 117957 (2024)
8. Najafi, N., Paulsen, U.S.: Operational modal analysis on a VAWT in a large wind tunnel using stereo vision technique. Energy **125**, 405–416 (2017)
9. Peeters, B., De Roeck, G.: Reference-based stochastic subspace identification for output-only modal analysis. Mechanical Systems and Signal Processing **13**(6), 855–878 (1999)
10. Schreier, H., Orteu, J.J., Sutton, M.A.: Image Correlation for Shape, Motion and Deformation Measurements. Springer US, Boston, MA (2009)
11. Stone, H., Orchard, M., Chang, E.C., Martucci, S.: A fast direct Fourier-based algorithm for subpixel registration of images. IEEE Transactions on Geoscience and Remote Sensing **39**(10), 2235–2243 (2001)
12. The MathWorks Inc.: vision.PointTracker (2023), https://www.mathworks.com/help/vision/ref/vision.pointtracker-system-object.html
13. Wang, Y.Q., Sutton, M.A., Bruck, H.A., Schreier, H.W.: Quantitative error assessment in pattern matching: Effects of intensity pattern noise, interpolation, strain and image contrast on motion measurements. Strain **45**(2), 160–178 (2009)
14. Wang, Y., Egner, F.S., Willems, T., Kirchner, M., Desmet, W.: Camera-based experimental modal analysis with impact excitation: Reaching high frequencies thanks to one accelerometer and random sampling in time. Mechanical Systems and Signal Processing **170**, 108879 (2022)
15. Xin, C., Wang, C., Xu, Z., Qin, M., He, M.: Marker-free vision-based method for vibration measurements of RC structure under seismic vibration. Earthquake Engineering & Structural Dynamics **51**(8), 1773–1793 (2022)
16. Ye, X.W., Yi, T.H., Dong, C.Z., Liu, T.: Vision-based structural displacement measurement: System performance evaluation and influence factor analysis. Measurement **88**, 372–384 (2016)
17. Yu, S., Su, Z., Zhang, J.: Robust optical displacement measurement of bridge structures in complex environments. ISPRS Journal of Photogrammetry and Remote Sensing **192**, 395–408 (2022)
18. Zhang, Y., Cao, G.: Transverse and longitudinal vibration measurement of vertical rope of hoisting system based on machine vision. IEEE Transactions on Instrumentation and Measurement **71**, 1–11 (2022)