Operational modal analysis of a cantilever beam by phase-based sub-pixel motion estimated from experimental videos

Cédric Marinel, Olivier Losson, Benjamin Mathon, Jean Le Besnerais, Ludovic Macaire

Operational modal analysis of a cantilever beam by phase-based motion estimated from experimental videos

C. Marinel[†⋆], O. Losson[†], B. Mathon[†], J. Le Besnerais[⋆], and L. Macaire[†]

[†]Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France
[⋆]EOMYS Engineering, Villeneuve d'Ascq, France

ABSTRACT

Vibration study, also called modal analysis, plays an important role in the structural health monitoring of mechanical structures. During the last decade, video-based modal analysis methods have emerged to provide dense vibration estimation using each pixel as a contactless sensor. Dense subpixel motion is estimated and is then processed by a modal analysis algorithm to extract the modal basis composed of natural frequencies, damping ratios, and mode shapes. This paper introduces a new single-subband phase-based method for subpixel motion estimation. It is compared with state-of-the-art motion estimation method on synthetic and experimental videos of a cantilever beam.

Keywords: Operational Modal Analysis, Vibration, Structural Health Monitoring, Subpixel motion estimation, Phase-based motion estimation

1. INTRODUCTION

Monitoring mechanical properties of large mechanical structures such as bridges, building or wind turbines is important to detect early stage failures. Operational modal analysis (OMA) estimates vibration properties, such as resonance frequencies, from motion measurements. Change of resonance frequencies in time may be a good indicator of failure.[1] OMA generally analyzes signals provided by contact sensors, such as accelerometers or laser vibrometers. However, those are expensive and may be hard to set up. Laser vibrometers enable remote sensing, but only measure one-directional motion point by point. Because the force applied to the structure may be different as each point motion is measured, another sensor is required to measure a reference fixed point during the whole acquisition.

Taking advantage of fast video camera development and of works about subpixel motion estimation, several video-based OMA methods have emerged in the last decade. These methods combine motion estimation methods with traditional modal analysis algorithms to allow for a cheap remote sensing. Among classical methods that estimate subpixel motion from videos, phase-based ones provide a dense motion estimation with no need of speckle patterns or targets mounted on the structure.[2] This strategy decomposes spatial frequencies within each frame into complex subbands by applying Gabor filters with localized frequencies. Chen *et al.* analyze on vertical and horizontal subbands to estimate 2D motion.[3] They assume that motion can be captured using a single arbitrarily chosen scale. However, spatial frequencies, and thus the optimal scale, may depend on the spatial resolution. Yang *et al.* only focus on horizontal motion and apply a single-scale method to estimate the motion on each horizontal subband.[4] Wadhwa *et al.* propose a multi-subband approach thanks to a weighted least-square estimator to compute 2D motion.[5] However, weighted least-square estimation is time-consuming.

To reduce the processing time, we propose to select the relevant subband at each pixel in order to estimate motion. In this paper, we detail this method, denoted as single-subband approach. We compare its performance with that of the multi-subband method on both synthetic and experimental videos of a cantilever beam that moves along horizontal and vertical directions.

2. PHASE-BASED SUBPIXEL MOTION ESTIMATION

Let $\boldsymbol{\delta}[\boldsymbol{p};t] = (\delta^h[\boldsymbol{p};t], \delta^v[\boldsymbol{p};t])^\intercal \in \mathbb{R}^2$ be the motion field, defined by its coordinates along horizontal and vertical directions, at pixel $\boldsymbol{p}$ in frame $t \in [\![0, N_t - 1]\!]$. Assuming illumination is spatially and spectrally constant over time, the intensity $I[\boldsymbol{p};t]$ at the pixel $\boldsymbol{p}$ associated to a given surface element can be considered as constant:

$$I[\boldsymbol{p};0] = I[\boldsymbol{p} + \boldsymbol{\delta}[p;t];t]. \tag{1}$$

Furthermore, the intensity $I[\boldsymbol{p};t]$ is assumed to be a combination of local cosine functions that depend on local spatial frequencies. The spatial frequency domain may be divided into subbands, each being centered at spatial frequency $\boldsymbol{\omega}_{r,\theta} = (\omega^h_{r,\theta}, \omega^v_{r,\theta})^\intercal = (\omega_r \cos(\theta), \omega_r \sin(\theta))^\intercal$. These subbands correspond to different scales $r \in [\![1, N_r]\!]$ and orientations $\theta = 0, \pi/N_\theta, \dots, (N_\theta - 1)\pi/N_\theta$, where N_r and N_θ are the chosen numbers of scales and orientations.

Frame spatial frequencies are decomposed into subbands using Gabor filters $G_{r,\theta}$, whose band-pass filter around central frequency $\boldsymbol{\omega}_{r,\theta}$ is characterized by a standard deviation $\sigma_{r,\theta}$. For each subband (r,θ), the complex filter response $S_{r,\theta}[\boldsymbol{p};t] = G_{r,\theta}[\boldsymbol{p}] * I[\boldsymbol{p};t]$ is expressed as:

$$S_{r,\theta}[\boldsymbol{p};t] = \rho_{r,\theta}[\boldsymbol{p};t] \cdot \exp\left(j \cdot (\tilde{\boldsymbol{\omega}}_{r,\theta}[\boldsymbol{p}]^\intercal \cdot (\boldsymbol{p} - \boldsymbol{\delta}[\boldsymbol{p};t]))\right), \tag{2}$$

where $\tilde{\boldsymbol{\omega}}_{r,\theta}[\boldsymbol{p}]$ is the local spatial frequency in the subband centered at $\boldsymbol{\omega}_{r,\theta}$ such that $|\tilde{\boldsymbol{\omega}}_{r,\theta}[\boldsymbol{p}] - \boldsymbol{\omega}_{r,\theta}[p]| < \sigma_{r,\theta}$, $\rho_{r,\theta}[\boldsymbol{p};t] = |S_{r,\theta}[\boldsymbol{p};t]|$ is the filter response magnitude, and its phase is:

$$\varphi_{r,\theta}[\boldsymbol{p};t] = \arctan\left(\mathfrak{Im}(S_{r,\theta}[\boldsymbol{p};t])/\mathfrak{Re}\left(S_{r,\theta}[\boldsymbol{p};t]\right)\right) = \tilde{\boldsymbol{\omega}}_{r,\theta}[\boldsymbol{p}]^\intercal \cdot (\boldsymbol{p} - \boldsymbol{\delta}[\boldsymbol{p};t]). \tag{3}$$

Assuming there is no motion in the first frame (*i.e.*, $\boldsymbol{\delta}[\boldsymbol{p};0] = 0$ for all $\boldsymbol{p}$), Eq. (3) yields the following for each subband (r,θ), pixel $\boldsymbol{p}$, and frame t:

$$\varphi_{r,\theta}[\boldsymbol{p};0] - \varphi_{r,\theta}[\boldsymbol{p};t] = \tilde{\boldsymbol{\omega}}_{r,\theta}[\boldsymbol{p}]^\intercal \cdot \boldsymbol{\delta}[\boldsymbol{p};t]. \tag{4}$$

This equation is important because it links the motion $\boldsymbol{\delta}$ with the local spatial frequency $\tilde{\boldsymbol{\omega}}_{r,\theta}$ and the phase $\varphi_{r,\theta}$. The phase is computed from the complex response $S_{r,\theta}$ at each pixel within each frame.

2.1 Multi-subband 2D motion estimation

Wadhwa *et al.* propose to take account of the phase and local frequency computed in every subband, and compute a weighted least-square estimator of the 2D motion.[5] Because the response phase $\varphi_{r,\theta}$ is only meaningful if the associated amplitude is high, the authors use its square $\rho^2_{r,\theta}$ as weight. They also assume that motion is locally smooth, and add a spatial consistency constraint via a Gaussian kernel $\mathcal{G}$ (with 3 px standard deviation and 19×19 px support W). The local frequency $\tilde{\boldsymbol{\omega}}_{r,\theta}$ at each pixel $\boldsymbol{p}$ is approximated by the Gabor central frequency $\boldsymbol{\omega}_{r,\theta}$. Motion is then determined by a weighted least-square estimator as:

$$\hat{\boldsymbol{\delta}}[\boldsymbol{p};t] = \underset{\boldsymbol{\delta}[\boldsymbol{p};t]}{\arg\min} \sum_r \sum_\theta \sum_{\boldsymbol{q} \in W[\boldsymbol{p}]} \mathcal{G}[\boldsymbol{q}] \cdot \rho^2_{r,\theta}[\boldsymbol{p}+\boldsymbol{q};t] \cdot \left[\boldsymbol{\omega}^\intercal_{r,\theta} \cdot \boldsymbol{\delta}[\boldsymbol{p};t] - \left(\varphi_{r,\theta}[\boldsymbol{p}+\boldsymbol{q};0] - \varphi_{r,\theta}[\boldsymbol{p}+\boldsymbol{q};t]\right)\right]^2. \tag{5}$$

Before solving the weighted least-square problem, the phase $\varphi_{r,\theta}$ is temporally unwrapped. As this approach attempts to minimize a square distance between $N_r N_\theta$ terms at each pixel within each frame, it is time-consuming.

2.2 Single-subband 2D motion estimation

To reduce the processing time, we propose to select the relevant subband at each pixel. For each pixel $\boldsymbol{p}$, the retained subband $(r_{\boldsymbol{p}}, \theta_{\boldsymbol{p}})$ is that with maximum amplitude within the first frame:

$$(r_{\boldsymbol{p}}, \theta_{\boldsymbol{p}}) = \underset{r,\theta}{\operatorname{argmax}}\ \rho_{r,\theta}[\boldsymbol{p},0]. \tag{6}$$

Denoting $\tilde{\boldsymbol{\omega}}_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p}] = \left|\tilde{\boldsymbol{\omega}}_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p}]\right| \mathbf{u}$, with $\mathbf{u} = \left(\cos\left(\angle_{\tilde{\boldsymbol{\omega}}_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p}]}\right), \sin\left(\angle_{\tilde{\boldsymbol{\omega}}_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p}]}\right)\right)^{\intercal}$, and assuming that motion $\boldsymbol{\delta}[\boldsymbol{p};t]$ is collinear with $\tilde{\boldsymbol{\omega}}_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p}]$, *i.e.*, $\boldsymbol{\delta}[\boldsymbol{p};t] = |\boldsymbol{\delta}[\boldsymbol{p};t]|\mathbf{u}$, Eq. (4) becomes:

$$\varphi_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p};0] - \varphi_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p};t] = |\tilde{\boldsymbol{\omega}}_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p}]| \cdot |\boldsymbol{\delta}[\boldsymbol{p};t]|. \tag{7}$$

Motion can thus be estimated using the selected subband $(r_{\boldsymbol{p}}, \theta_{\boldsymbol{p}})$ as:

$$\hat{\boldsymbol{\delta}}[\boldsymbol{p};t] = \frac{\varphi_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p};0] - \varphi_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p};t]}{|\tilde{\boldsymbol{\omega}}_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p}]|} \cdot \mathbf{u}. \tag{8}$$

At each pixel $\boldsymbol{p}$, the local spatial frequency within the selected subband $(r_{\boldsymbol{p}}, \theta_{\boldsymbol{p}})$ must be determined in order to estimate motion according to Eq. (8). For this purpose, we assume that motion and local frequency in the selected subband are locally constant within the frame, *i.e.*, $|\nabla_{\boldsymbol{p}}\, \boldsymbol{\delta}[\boldsymbol{p};t]| = 0$ and $\left|\nabla_{\boldsymbol{p}}\, \tilde{\boldsymbol{\omega}}_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p}]\right| = 0$, where $\nabla_{\boldsymbol{p}}$ is the spatial gradient computed at $\boldsymbol{p}$. In the selected subband $(r_{\boldsymbol{p}}, \theta_{\boldsymbol{p}})$, the local spatial frequency is then determined as $\tilde{\boldsymbol{\omega}}_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p}] = \nabla_{\boldsymbol{p}}\, \varphi_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p};0]$ by deriving Eq. (3) with respect to $\boldsymbol{p}$ within the first frame. To compute the phase gradient, we follow the approach of Fleet *et al.*[6] that avoids phase wrapping and discontinuity problems encountered with Sobel derivation, which provides:

$$\tilde{\boldsymbol{\omega}}_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p}] = \nabla_{\boldsymbol{p}}\, \varphi_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p};0] = \frac{\mathfrak{Im}\left(S^{*}_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p};0] \cdot \nabla_{\boldsymbol{p}}\, S_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p};0]\right)}{\left|\rho_{r_{\boldsymbol{p}},\theta_{\boldsymbol{p}}}[\boldsymbol{p};0]\right|^{2}}. \tag{9}$$

3. MODAL ANALYSIS

From the estimated dense motion, the goal of OMA is to estimate a modal basis $\{(f_i, \zeta_i, \phi_i)\}_{i=1}^{N_m}$ of N_m modes, each of which with resonance frequency f_i (Hz), damping ratio ζ_i, and mode shape ϕ_i. In this study, we use the covariance-driven stochastic subspace identification (SSI) method,[7] that provides a high parameter estimation accuracy for a short computation time,[8] as well as a stabilization diagram to discard spurious modes.

3.1 Covariance-driven stochastic subspace identification

The free motion equation for a system with N degrees of freedom is formulated as a discrete-time state space model:

$$\begin{aligned} \boldsymbol{x}[t+1] &= \boldsymbol{A}\boldsymbol{x}[t] + \boldsymbol{w}[t] \\ \boldsymbol{y}[t] &= \boldsymbol{C}\boldsymbol{x}[t] + \boldsymbol{v}[t]\ , \end{aligned} \tag{10}$$

where $\boldsymbol{x}[t] \in \mathbb{R}^{2N}$ is the state vector, $\boldsymbol{y}[t] \in \mathbb{R}^{2N_p}$ the observation vector for N_p pixels of interest, $\boldsymbol{A} \in \mathbb{R}^{2N\times 2N}$ the state-space matrix, $\boldsymbol{C} \in \mathbb{R}^{2N_p\times 2N}$ the observation matrix, and $\boldsymbol{w}[t] \in \mathbb{R}^{2N}$ and $\boldsymbol{v}[t] \in \mathbb{R}^{2N_p}$ the observation and input noise vectors. The goal of SSI[7] is to get estimates $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{C}}$ of these matrices only from observations $\{\boldsymbol{y}[t]\}_{t=0}^{N_t-1}$ (here, $\hat{\boldsymbol{\delta}}[:,t]$) to obtain the modes. A block Toeplitz matrix composed of observation covariance matrices with different time shifts is constructed. The singular value decomposition of the Toeplitz matrix provides $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{C}}$. Then, from the eigenvalue decomposition $\hat{\boldsymbol{A}} = \boldsymbol{\Psi}\boldsymbol{S}\boldsymbol{\Psi}^{-1}$ with $\boldsymbol{S} = \mathrm{diag}(\lambda_i), i \in [\![1, 2N]\!]$, the modal basis is computed for a given model order N as:

$$f_i^N = \frac{1}{2\pi}\left|\frac{\log \lambda_i}{\Delta t}\right|,\ \zeta_i^N = \frac{\mathfrak{Re}(\lambda_i)}{|\lambda_i|},\ \text{and}\ \boldsymbol{\Phi}^N = \hat{\boldsymbol{C}}\boldsymbol{\Psi}, \tag{11}$$

where Δt is the time step between two successive frames. Mode shapes ϕ_i^N are the columns of $\boldsymbol{\Phi}^N$. Note that the $2N$ modes come as N complex conjugate pairs, hence N positive frequencies are kept.

3.2 Stabilization diagram

In practice, the number N_m of modes is unknown. Therefore, a stabilization diagram is used to plot the poles obtained from a modal analysis method with different model orders.[9] Irrelevant models produce spurious modes

that can be discarded by a stability analysis. Indeed, physical poles tend to be stable whereas spurious ones tend to be unstable (in frequency, damping, and/or mode shape).

Let p_i^N, $i \in [\![1, N]\!]$, be a pole of order N with natural frequency f_i^N, damping ratio ζ_i^N, and mode shape $\boldsymbol{\phi}_i^N$. We consider it as stable if the following predicate holds:

$$S\left(p_i^N\right) = \exists p_j^{N-1} \left[\left(\frac{|f_i^N - f_j^{N-1}|}{f_j^{N-1}} < 0.01 \right) \wedge \left(\frac{|\zeta_i^N - \zeta_j^{N-1}|}{\zeta_j^{N-1}} < 0.05 \right) \wedge \left(\mathrm{MAC}(\boldsymbol{\phi}_i^N, \boldsymbol{\phi}_j^{N-1}) > 0.98 \right) \right], \tag{12}$$

where $\mathrm{MAC}(\boldsymbol{\phi}_a, \boldsymbol{\phi}_b) = \dfrac{(\boldsymbol{\phi}_a^\intercal \boldsymbol{\phi}_b)^2}{(\boldsymbol{\phi}_a^\intercal \boldsymbol{\phi}_a)(\boldsymbol{\phi}_b^\intercal \boldsymbol{\phi}_b)}$ is the modal assurance criterion between any two mode shapes $\boldsymbol{\phi}_a$ and $\boldsymbol{\phi}_b$. The key idea is that spurious modes occur randomly and are not stable for two consecutive model orders.

However, when motion measurements are noisy, some spurious modes can be confused with stable ones on the diagram. To discard them, the first complex mode indication function (CMIF), whose peaks indicate the presence of a mode at the associated frequency, is also plotted on the diagram.[10]

4. EXPERIMENTS

To compare the methods presented in Sec. 2, we here analyze two videos of a cantilever beam: first, a synthetic noise-free video generated using the Euler–Bernoulli beam model and, second, an experimental video recorded in controlled conditions by a TIS DMK 33UX287 camera.

4.1 Synthetic video of two-directional motion

4.1.1 Setup

We generate a synthetic video of a cantilever beam using the Euler–Bernoulli model with the following physical parameters: beam length $L = 0.9$ m, Young modulus $E = 210 \cdot 10^9$ Pa, area moment of inertia $J = 5.4 \cdot 10^{-10}$ m^4, and mass per unit length $\mu = 1.413$ $\mathrm{kg{\cdot}m^{-1}}$. The center line of the beam is defined in the scene coordinate system by the point set $\{(g(z,t), z; t) \in \mathbb{R} \times [0, L] \times [0, (N_t - 1)\Delta t]\}$, where $g(z,t) = \sum_{m=1}^{N_m} \phi_m(z)\ q_m(t)$ is the Euler–Bernoulli equation solution, in which a mode shape ϕ_m and modal coordinate q_m respectively correspond to space and time information.

An input force $\gamma = 0.32\,\mathrm{N}$ at $t = 0$ and $z = L$ simulates a hammer impact at the free end of the beam in its normal direction. Four modes are defined in the model ($N_m = 4$). Theoretical natural frequencies $\{f_m\}_{m=1}^4$ are computed from the Euler–Bernoulli equation, and damping ratios are set arbitrarily. Their values are listed in Tab. 1.

Mode m	1	2	3	4
f_m (Hz)	6.19	38.79	108.60	212.82
ζ_m (%)	0.11	1.13	0.29	0.13

Table 1. Beam theoretical natural frequencies and damping ratios.

Videos last 1 s at 436 fps (default frame rate of TIS DMK 33UX287 camera), with a frame size set to 1000×1000 px and pixel resolution of $1.289\,\mathrm{mm{\cdot}px^{-1}}$. Each pixel intensity depends on the area of the intersection between this pixel in the image plane and the projected beam using a pinhole camera model without optical distortion. Gray levels are then encoded on 8 bits. As the beam motion is mainly in its normal direction, each frame is rotated by an angle of 87.75° to study two-directional subpixel motion. Motion is estimated at each pixel, and N_p pixels of interest are then extracted from a line between the beam root and free end to perform modal analysis (see Fig. 1).

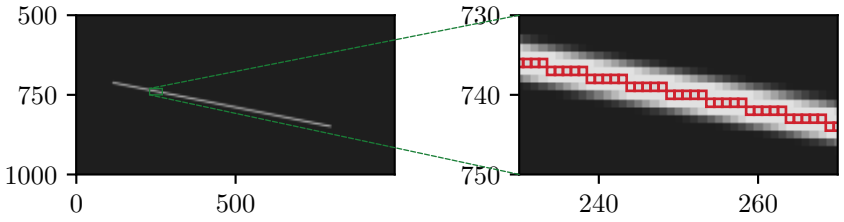

Figure 1. Frame from synthetic video and pixels of interest (in red).

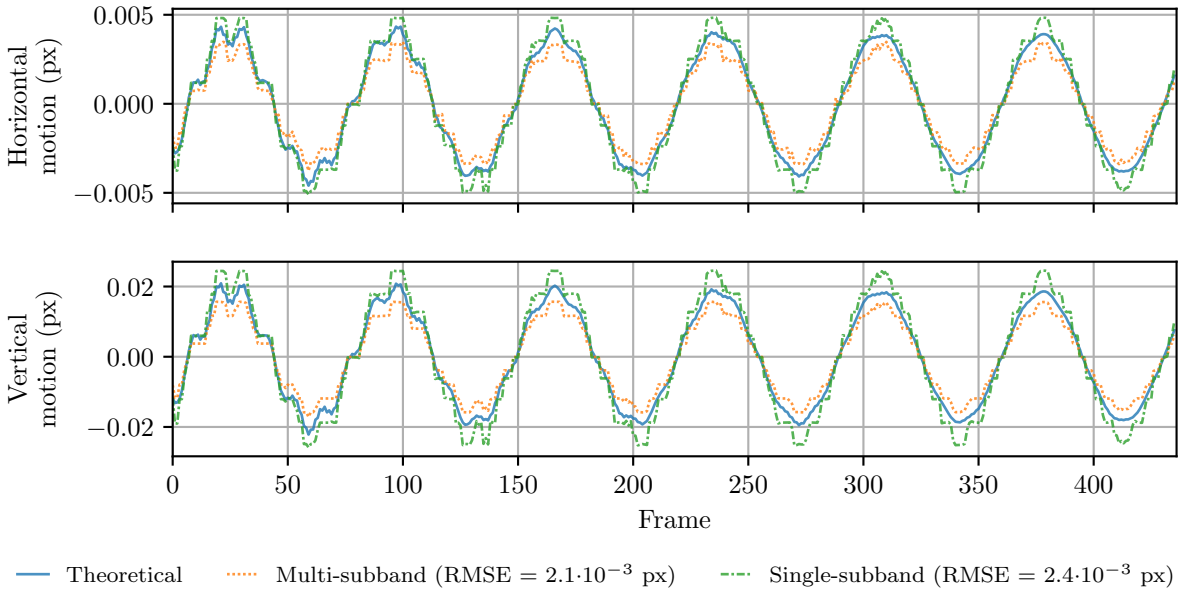

Figure 2. Estimated and theoretical motion in the middle of the beam. RMSE is the root-mean-square error.

4.1.2 Results

For both methods, the numbers of scales and orientations are set to $N_r = 5$ and $N_\theta = 8$. Theoretical and estimated motions along horizontal and vertical directions are shown in Fig. 2. The multi-subband estimator slightly underestimates motion while the single-subband method overestimates it, but both methods give satisfying subpixel motion estimation. The stabilization diagrams are constructed with model orders $N \in [\![4, 30]\!]$ and are displayed in Fig. 3. We can see that the diagrams are very similar, and that multiple stable poles are present at the theoretical natural frequencies. To get quantitative results (see Tab. 2), we select the stable poles close (by 1%) to a theoretical frequency, then we compare their average damping ratio and mode shape (using the MAC) to the theoretical ones. The number of stable poles and damping ratios are comparable for both methods. The single-subband method is less effective than the multi-subband one in terms of mode shape error, but is three times quicker to process (2214 s vs. 6575 s).

Criterion	Number of stable poles				Damping ratio absolute error				Mode shape error (1-MAC)			
Mode m	1	2	3	4	1	2	3	4	1	2	3	4
Multi-subband	**17**	**26**	21	6	0.202	**0.013**	**0.002**	**0.046**	**0.000**	**0.009**	**0.077**	**0.371**
Single-subband	15	**26**	**24**	**10**	**0.173**	0.012	0.008	0.056	0.001	0.012	0.086	0.449

Table 2. Comparison of OMA results of both motion estimation methods with theoretical ones on synthetic video.

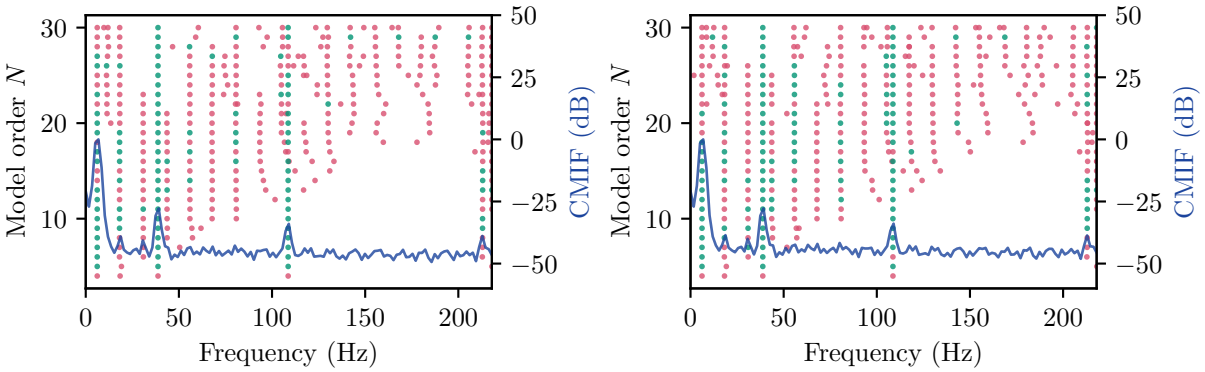

Figure 3. Stabilization diagrams obtained for multi-subband (left) and single-subband (right) methods from the analysis of synthetic video. Stable poles are in green and unstable ones in red.

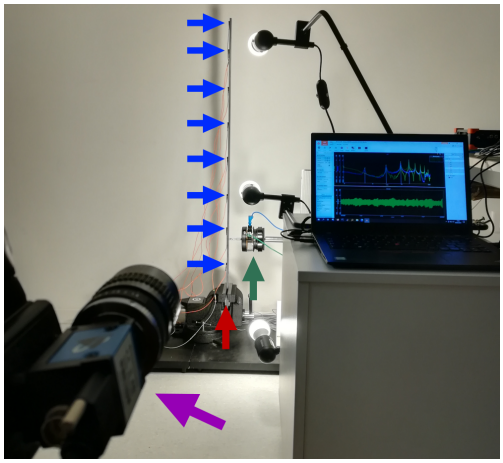

Figure 4. Experimental setup with camera in purple, beam in red, accelerometers in blue, and vibration shaker in green.

4.2 Experimental video

4.2.1 Setup

To compare the methods on experimental videos, we clamp a steel beam of $930 \times 30 \times 6\,\text{mm}^3$ in a vice. We place eight single-axis accelerometers at evenly-spaced locations on the beam to get its ground truth in natural frequencies and damping ratios. The beam is excited by a vibration shaker with a random white noise between 0 and 250 Hz. The experimental setup is shown in Fig. 4. The modes listed in Tab. 3 are estimated using the covariance-driven SSI algorithm. As the frequency range of accelerometers is above 10 Hz, the first mode is not retrieved in practice. Nevertheless, the marked peak at 5.31 Hz on the acceleration power spectral densities provides a coarse estimation of the first natural frequency.

Mode m	1	2	3	4
f_m (Hz)	5.31	30.69	96.49	186.72
ζ_m (%)	not estimated	1.16	1.30	0.25

Table 3. Beam natural frequencies and damping ratios estimated by accelerometers.

The TIS DMK 33UX287 camera with a 8 mm focal length lens is installed in front of the beam and tilted by an angle of 9° in order to be close to the synthetic video case. The frame rate is set to 500 fps, the acquisition time to 8 s, and a ROI of size of 720×140 px is used. The motion estimation parameters are the same as for synthetic video ($N_r = 5$, $N_\theta = 8$). A frame with the pixels of interest is shown in Fig. 5.

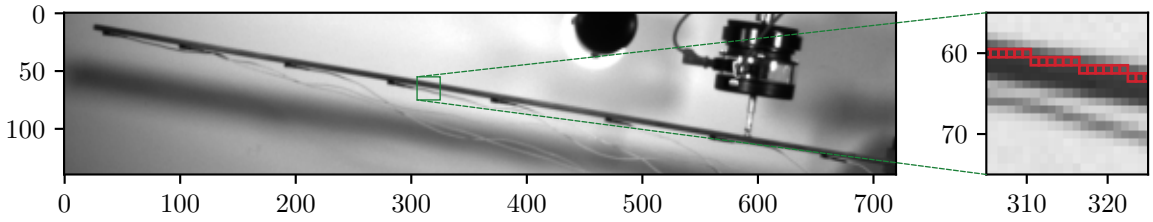

Figure 5. Frame from experimental video and pixels of interest (in red).

A stabilization diagram is built, then damping ratios and mode shapes of stable poles (close to natural frequencies estimated with accelerometers) are compared to ground truth damping ratios and theoretical mode shapes.

4.2.2 Results

The stabilization diagrams obtained for both methods are displayed in Fig. 6 and the quantitative results are gathered in Tab. 4. The four modes are identifiable on both stabilization diagrams. However, a spurious mode close to 200 Hz, possibly caused by video noise that corrupts motion estimation, is also present on both diagrams. The single-subband method gives competitive results compared to the multi-subband one, for a reduced motion estimation time (2834 s vs. 6387 s).

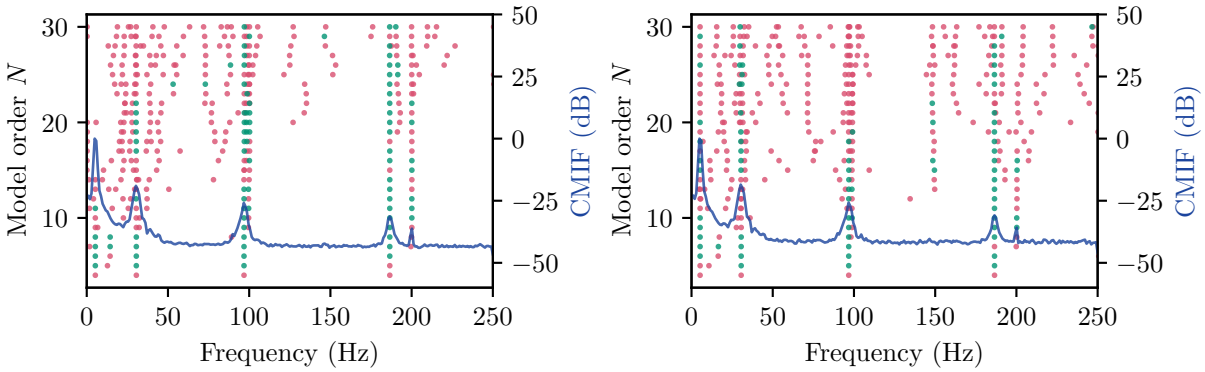

Figure 6. Stabilization diagram for multi-subband (left) and single-subband (right) methods on experimental video.

Criterion	Number of stable poles				Damping ratio absolute error				Mode shape error (1-MAC)			
Mode m	1	2	3	4	1	2	3	4	1	2	3	4
Multi-subband	5	**10**	**19**	**22**	–	**1.288**	0.142	0.333	**0.005**	0.288	0.037	0.115
Single-subband	**11**	7	12	21	–	1.545	**0.068**	**0.296**	0.007	**0.097**	**0.034**	**0.078**

Table 4. Comparison of OMA results of both motion estimation methods with accelerometer ones on experimental video.

5. CONCLUSION

In this paper, we introduce a new single-subband phase-based method to estimate the 2D motion of a cantilever beam from subpixel video analysis. Spatial frequencies of each frame are decomposed into subbands with complex Gabor filters. The proposed method uses the phase of the complex subband with the highest amplitude at each pixel. We also consider a multi-subband method based on a least square estimator merging all subband phases.

The two methods are compared on synthetic and experimental videos of a cantilever beam. The synthetic video is generated using the Euler–Bernoulli equation of the beam excited by a hammer impact. The experimental video is acquired in controlled lighting conditions and the beam is excited by a vibration shaker with white noise. Motion is estimated in every frame and is then extracted on a line of beam pixels. The modal basis, composed of

natural frequency, damping ratio, and mode shape, is then computed from motion estimations. The results on synthetic and experimental videos are respectively compared to the theoretical modal basis of an Euler–Bernoulli cantilever beam model and to the modal basis estimated from accelerometer measurements. Both methods provide good motion estimations on synthetic video, with an RMSE to theoretical motion below $3 \cdot 10^{-3}$ px. For both videos, the single-subband method gives comparable modal basis estimates w.r.t. to the multi-subband method. Furthermore, the single-subband method reduces the computation time by a factor between 2 and 3.

Further tests on more complex structures should be performed to confirm the results on experimental videos. Moreover, experiments on videos acquired in outdoor should be performed to generalize this work to uncontrolled video acquisition conditions.

REFERENCES

[1] Salawu, O., "Detection of structural damage through changes in frequency: a review," *Engineering Structures* **19**, 718–723 (Sept. 1997).

[2] Chou, J.-Y. and Chang, C.-M., "Image motion extraction of structures using computer vision techniques: A comparative study," *Sensors* **21** (Sept. 2021). paper 6248.

[3] Chen, J. G., Wadhwa, N., Cha, Y.-J., Durand, F., Freeman, W. T., and Buyukozturk, O., "Modal identification of simple structures with high-speed video using motion magnification," *Journal of Sound and Vibration* **345**, 58–71 (June 2015).

[4] Yang, Y., Dorn, C., Mancini, T., Talken, Z., Kenyon, G., Farrar, C., and 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 (Feb. 2017).

[5] Wadhwa, N., Chen, J. G., Sellon, J. B., Wei, D., Rubinstein, M., Ghaffari, R., Freeman, D. M., Büyüköztürk, O., Wang, P., Sun, S., Kang, S. H., Bertoldi, K., Durand, F., and Freeman, W. T., "Motion microscopy for visualizing and quantifying small motions," *Proceedings of the National Academy of Sciences* **114**, 11639–11644 (Oct. 2017).

[6] Fleet, D. J. and Jepson, A. D., "Computation of component image velocity from local phase information," *International Journal of Computer Vision* **5**, 77–104 (Aug. 1990).

[7] Peeters, B. and de Roeck, G., "Reference-based stochastic subspace identification for output-only modal analysis," *Mechanical Systems and Signal Processing* **13**, 855–878 (Nov. 1999).

[8] Zahid, F. B., Ong, Z. C., and Khoo, S. Y., "A review of operational modal analysis techniques for in-service modal identification," *J Braz. Soc. Mech. Sci. Eng.* **42**, 398 (Aug. 2020).

[9] Van der Auweraer, H. and Peeters, B., "Discriminating physical poles from mathematical poles in high order systems: use and automation of the stabilization diagram," in [*Procs. 21st IEEE Instrumentation and Measurement Technology Conference (IMTC'04)*], **3**, 2193–2198 (May 2004).

[10] Shih, C., Tsuei, Y., Allemang, R., and Brown, D., "Complex mode indication function and its applications to spatial domain parameter estimation," *Mechanical Systems and Signal Processing* **2**, 367–377 (Oct. 1988).