

Single-pixel optical-flow-based experimental modal analysis

I. Tomac^{a,b,*}, J. Slavič^{a,**}, D. Gorjup^a

^a*University of Ljubljana, Faculty of Mechanical Engineering, Aškerčeva cesta 6, SI-1000 Ljubljana, Slovenia*

^b*University of Split, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, Ruđera Boškovića 32, HR-21000 Split, Croatia*

Cite as:

Ivan Tomac, Janko Slavič, Domen Gorjup; Single-pixel optical-flow-based experimental modal analysis, Mechanical Systems and Signal Processing, Volume 202, November 2023, 110686

DOI: 10.1016/j.ymssp.2023.110686

Abstract

Modal analysis using structural responses identified from high-speed cameras is a challenging task. The problem is that the measured displacements are relatively small (typically deep in the sub-pixel range) and submerged in noise due to the low dynamic range of the camera sensor. A typical approach to determine structural responses from high-speed camera data is the digital image correlation (DIC) method, a general, computationally intensive method for identifying displacements. Without knowing the assumptions of the modal analysis, DIC identifies the displacement in the time domain by minimising the difference between two consecutive regions of interest (ROIs). Optical flow is a method based on the change in intensity in a given pixel due to the change in reflection from a moving surface. The displacement is identified from the change in intensity and the spatial gradient of the intensity of the surface. For small, sub-pixel movements, the relationship between intensity change and displacement is linear, which opens up the possibility of performing the modal analysis directly on the pixel intensity measured by the camera. This research applies the recently introduced Morlet-wave modal method and introduces an experimental modal analysis based on a single pixel with optical flow directly from the pixel intensities and the spatial gradient of the intensity. Furthermore, it is shown that the natural frequencies and damping ratios do not require the spatial gradient. The introduced method was successfully applied to the experimental test case where a pixel-based, full-field modal analysis was performed. The influence of averaging the results from multiple pixels in the modal domain is investigated. Modal identification is compared with the results obtained from the displacements identified with a digital image correlation (DIC) method. The introduced direct pixel-based modal analysis provides a robust and numerically efficient way to a full-field modal analysis.

Keywords: Morlet-wave, single pixel, modal identification, full-field

1. Introduction

The main advantage of using high-speed cameras in structural dynamics analyses is the non-contact and relatively simple measurement of the full-field structural response [1], typically providing several thousand measurement points. The advantage of such an over-determination of the response has been noted in several areas of structural dynamics, such as structural health monitoring [2], damage localisation [3], model updating [4], measurement [5] and reconstruction [6] of the sound-radiation field, and system identification [7]. In most of these areas, the analysis is made in the first step with displacement responses obtained by processing high-speed video images using the digital image correlation (DIC) method [8]. DIC is based on an optimisation procedure performed on selected regions of interest (ROI) between two consecutive images [9]. This method is computationally intensive and time consuming. The second step is then to perform a modal identification using the selected method, which requires additional computational time [10, 11]. The identified displacements are typically in the subpixel range, where noise dominates due to the low dynamic range, which poses a challenge for the modal identification, especially at higher modes [12].

Modal identification from high-speed video data has seen significant research interest, for example, Huňady and Hagara [13] (2017) presented a method for full-field modal analysis based on the frequency-response functions (FRFs) of the full-field structural response obtained using the DIC method. The presented method for modal analysis is based on the singular-value decomposition (SVD) of FRFs to identify the mode shapes and decouple the system into equivalent single-degree-of-freedom (SDOF) FRFs to obtain Enhanced FRFs. These were obtained as a weighted average of all the FRFs using singular vectors as weighting functions to identify the natural frequencies and damping ratios with the Rational Fraction Polynomial [14] and Frequency-Domain Polynomial [15] methods. Yang et al. [16] (2017) performed a modal analysis of the responses obtained using the phase-based optical flow method [17]. The principal-component-analysis method [18] and the blind-source-separation technique [19] were used to reduce the dimensionality and isolate the vibration modes, and dense mode shapes were identified; natural frequencies and damping ratios were identified using the Hilbert-transform-based SDOF method [20]. Javh et al. [21] (2018) identified the mode shapes from responses obtained with a gradient-based optical flow up to 10 kHz using the LSFD method [22], but the system poles were identified with a high dynamic-range-sensor using the LSCF method [23]. Silva et al. [24] (2020) processed the pixel intensities with the non-negative matrix-factorisation algorithm [25] to generate the basis for the blind-source-separation technique [19] to separate vibrating modes and identify the mode shapes; the natural frequencies were estimated with a Fourier transform and the damping ratios were identified with a logarithmic decrement. Yang and Dorn [26] (2020) obtained the structural response using the phase-based optical flow method [16] and then

*Corresponding author

**Principal corresponding author

Email addresses: `itomac@fesb.hr` (I. Tomac), `janko.slavic@fs.uni-lj.si` (J. Slavič), `domen.gorjup@fs.uni-lj.si` (D. Gorjup)

performed mode separation and mode-shape identification using the affinity-propagation-cluster method [27]. Li et al. [28] (2022) developed an adaptive spatial filtering approach to the pixel intensities of sharp edges to fit the mode shapes and improve the modal displacements that are used for the identification of natural frequencies and damping ratios with the least-squares rational function method [29]. Wang et al. [30] (2022) used the Lucas-Kanade method of optical flow [31] on randomly sampled video data to perform modal identification on modes lying above the Shannon-Nyquist sampling limit; modal parameters were identified during the mode reconstruction optimisation process. By obtaining the full-field response of the structure with DIC, Ángel J. Molina-Viedma et al. [32] (2022) validated classical modal identification methods (PolyMAX [33] and LSCE [34]) also for the case of base excitation, where an input of the structure is defined as an controlled acceleration at the structure's fixation point.

This research is similar to Javh et al. [12] and takes each pixel as a sensor and then uses the Morlet-wave-modal method [35] to perform the modal identification. In this way the modal parameters are identified directly from the pixel intensities without the need to first retrieve the displacement response, which is due to the fact that, assuming a small displacement, the displacement of the object is linearly related to the change in intensity. The modal identification is demonstrated with an experimental laboratory test where the modal identification is performed up to 2.6 kHz.

In Sec. 2 the theoretical background of optical flow, the MW modal method and the discretisation of pixel intensities by the camera sensor are presented. The identification of modal parameters from the pixel intensities is introduced in Sec. 3. In Sec. 4, the single-pixel modal identification is demonstrated on a simple, numerical example, including a sensitivity analysis with respect to noise and low dynamic range. The MW modal method is applied to an experimental example in Sec. 5. The results based on the responses of the individual pixels are compared with the results obtained by averaging the identification results of more pixels in the modal domain, including checking with the results obtained from the accelerometer response. Sec. 6 concludes the manuscript.

2. Theoretical background

This section presents the theoretical concepts on which this research is based. The response of the mechanical system reproduced with light intensities based on the simplified optical flow method is presented [12]. The simple model of the camera sensor is presented to reveal the sources of error due to noise and quantisation in the generation of pixel intensities.

2.1. Optical flow

An object captured by the high-speed camera normally produces a sequence of greyscale images. Mathematically, a single image is a 2D matrix of the form $[M, N]$ (rows, columns) with pixel intensities (Fig. 1) representing the light reflected from the object. The image contains spatial information about the object, *e.g.*, its edges and textures. When the object moves, the movement is registered as a spatial change in pixel intensities. The movement is recorded with a sequence of images (frames), usually taken at fixed time intervals. Thus, a

video recording is a function of the pixel intensities in space and time: $I = f(x, y, t)$, where the pixel intensities are represented with Cartesian coordinate system (CS), as shown in Fig. 1. Due to the simplicity the pixel CS is aligned with the Cartesian CS. The movement

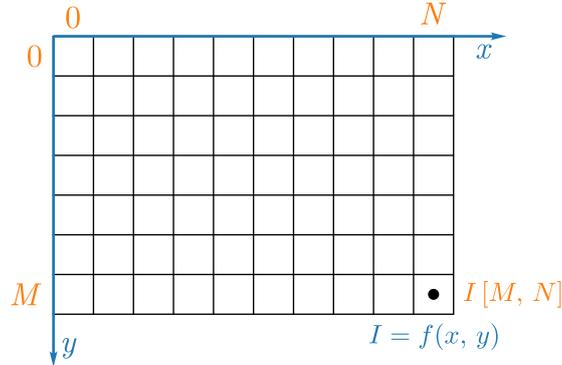


Figure 1: Coordinate systems of an computer image. Orange: pixel coord., blue: Cartesian coord.

is registered as a temporal change in the pixel intensities (the optical flow). In order to recognise the movement from the images, we must assume that the reflectivity of the observed object is constant and that the light source is constant; consequently, the recorded intensity is constant: $dI(x, y, t)/dt = 0$ [36]. For a line-scan camera (which records only one line in the x direction), the method of optical flow is shown in Fig. 2. The light reflected from the object entering the camera at time t has a (continuous) intensity $I(x, t)$; the camera at the observed pixel has an intensity $I(x_0, t)$. After Δt , the object has moved by Δx and the camera observes the pixel intensity $I(x_0 - \Delta x, t + \Delta t)$. For the assumption of small displacements

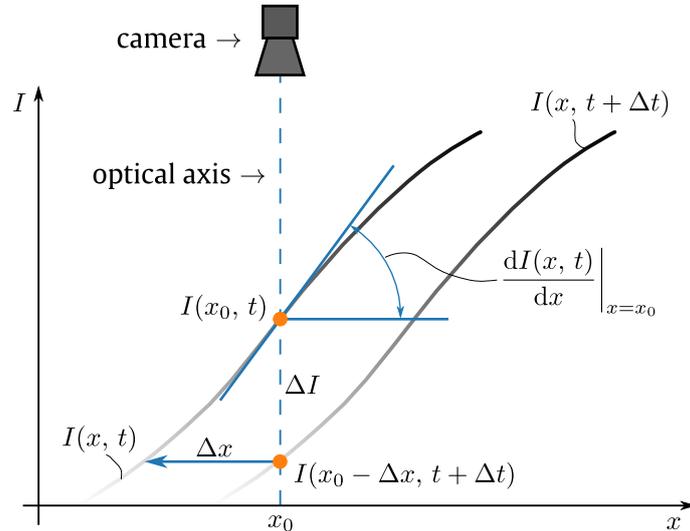


Figure 2: One-dimensional change of pixel intensities – $I(x, t)$, for continuous case.

(within the pixel), the change in intensity at the pixel $\Delta I = I(x_0 - \Delta x, t + \Delta t) - I(x_0, t)$

(Fig. 2) is associated with the gradient of intensity [12, 31]:

$$\frac{\Delta I}{\Delta x} \approx - \left. \frac{dI(x, t)}{dx} \right|_{x=x_0} \quad (1)$$

then the displacement based on a single-point measurement is [12]:

$$\Delta x \approx - \frac{\Delta I}{I'(x, t)} \quad (2)$$

where $I'(x, t) = dI/dx$ is the spatial gradient used to scale the change in pixel intensities into a displacement. To identify the displacements from a video recording, Eq. (2) must be iterated over the observed time interval for each pixel. If the displacement is small (*e.g.*, within one pixel), it is not necessary to calculate the spatial gradient (1) for each time step $j: t + j\Delta t$. Instead, a reference image $I_0(x)$ is obtained, from which the gradient image $I'_0(x)$ is then obtained. The reference image I_0 can be an image from any time step j , but it is best obtained by averaging the pixel intensities from several time steps to obtain an average image that is less affected by noise [12].

2.2. Light-intensity-based structural response

The displacement described with the light intensities in Fig. 2 shows a rigid-body motion; however, if the structure deforms (*e.g.*, oscillates harmonically), then the displacement of the object is a function of the spatial coordinate $q(x, t)$, which also affects the light reflection. Using the reference image I_0 to identify the displacements, Eq. (2) can be rewritten as follows:

$$q(x, t) = \frac{I_0(x) - I(x, t)}{I'_0(x)} \quad (3)$$

The extension to 2D optical flow requires the inclusion of the y direction in the Eq. (1), and the denominator in Eq. (3) must be replaced by $|\nabla I_0|$ [12]. To determine the gradient, information from the neighbouring pixels is needed, because the aperture problem cannot be solved in the single pixel, which is solved by using the smoothing condition [36]. However, the fact that the optical flow method identifies displacement in the direction of the gradient [36] allows us to perform research on the single-pixel basis. Displacement identification in 2D requires a high-gradient pattern that would allow camera to register in-plane motion *e.g.*, a speckle pattern. More general case where the plane is tilted for a general angle, adjacent pixels should be considered via averaging to reduce noise [12].

When the object is excited using an ideal impact, the light reflected from the object at the fixed point x_0 in Fig. 2 will harmonically oscillate around the static light intensity. Assuming a free, linear, viscously damped response, the oscillation of the reflected light can be represented by the sum of several SDOF systems, where the i -th mode is:

$$I_i(x, t) = A_{j,i}(x) e^{-\omega_{n,i} \delta_i t} \cos(\omega_{d,i} t - \phi_i(x)) \quad (4)$$

where $A_{j,i}$ is the amplitude of intensity and ϕ_i is the phase angle (both at location x and both depend on the initial conditions), δ_i is the damping ratio, $\omega_{n,i}$ is the undamped angular natural frequency and $\omega_{d,i} = \omega_{n,i} \sqrt{1 - \delta_i^2}$ is a damped angular natural frequency.

2.3. Morlet-wave modal method

The Morlet-wave (MW) modal identification method is based on the Morlet-wave integral to perform overdetermined modal identification from the free-response measurements in the time domain [35]. The Morlet-wave integral¹ is defined with:

$$\widetilde{W}(n, k, \omega) = \int_0^T f_m(t) \psi^*(n, k, \omega, t) dt, \quad (5)$$

where $f_m(t)$ is the measured signal, $\psi(t)$ is a Morlet-wave basis function and ψ^* is a complex conjugate function. The MW function is expressed with the parameters n , k , ω , where n is a time-spread parameter, k is the integer number of cycles of the wavelet function and ω is the frequency corresponding to the damped natural frequency ω_d of the selected mode [37]. The time length of the MW function is expressed by the number of cycles at the selected frequency $T = 2\pi k/\omega$. The theory of the continuous wavelet transform requires that the basic wavelet function must satisfy the admissibility condition [38]: $\int_{-\infty}^{\infty} \psi(t) dt = 0$. The basis function of the Morlet wave is defined on the finite interval $t \in [0, T]$; it is a symmetric function that translates for $T/2$ (for details see [37]) and for $k > 5$ and $n \geq 5$:

$$\int_0^T \psi(n, k, \omega, t) dt \approx 0. \quad (6)$$

The MW method inherits the linear property of the continuous wavelet transform $\mathcal{W}\{\cdot\}$, which allows separate analyses of the harmonic components of signals with multiple harmonic components:

$$\mathcal{W}_{n, k_i, \omega_i} \left\{ \sum_{i=1}^N a_i f_{m,i} \right\} = a_i \sum_{i=1}^N \mathcal{W}_{n, k_i, \omega_i} \{f_{m,i}\} \quad (7)$$

where a is a constant, the index i marks the frequency component and k_i is defined as:

$$k_i = \frac{T \omega_i}{2\pi} \quad (8)$$

If the analysed signals have closely spaced harmonic components, more attention is needed in the selection of the MW function parameters n and k [35].

To provide the basis for identifying the modal parameters, the MW integral of the free response (4) is derived analytically by paying attention to the theoretical limit, which is given by: $k \leq k_{\text{lim}} = n^2 \sqrt{1 - \delta^2} / (8\pi\delta)$ [35], resulting in the complex function:

$$W(n, k, \omega_d) \approx A \left(\frac{\pi}{2} \right)^{\frac{3}{4}} \sqrt{\frac{k}{n\omega_d}} e^{\frac{4\pi^2 k^2 \delta^2}{n^2(1-\delta^2)} - \frac{\pi \delta k}{\sqrt{1-\delta^2}}} e^{i(\pi k - \phi)} \varepsilon(n, k, \delta) \quad (9)$$

where $\varepsilon(n, k, \delta)$ is defined with the error functions [35]. The analytically obtained MW integral (9) is compared to the MW integral $\widetilde{W}(n, k, \omega)$ (5) obtained from the measured

¹In the literature [35, 37] the symbol I was used to denote the MW integral, but here it is used for the pixel intensities and W is used instead.

response. As described in the following, from the comparison the unknown modal parameters can be identified [35]. These modal parameters are identified in the order in which they are described.

The natural frequency is identified directly from the MW integral by numerically searching for the maximum of the MW integral around the initial natural frequency, obtained, for example, by picking peaks in the magnitude plot from the response in the frequency domain [35]:

$$\frac{\partial}{\partial \omega} \left| \widetilde{W}(n, k, \omega) \right| = 0 \quad (10)$$

The damping is then identified from the ratio between the absolute values of two MW integrals (5) obtained from the measured signal using different time-spread parameters n_1 , n_2 , for the given natural frequency ω_d and the predefined set of k parameters [35].

$$\widetilde{M}(n_1, n_2, k, \omega_d) = \frac{\left| \widetilde{W}(n_1, k, \omega_d) \right|}{\left| \widetilde{W}(n_2, k, \omega_d) \right|} \quad (11)$$

The same ratio is expressed analytically [35], where the natural frequency ω_d and the unknown amplitude A cancel out and an analytically expressed ratio is obtained $M(n_1, n_2, k, \delta)$, which allows the definition of a cost function $F_{\text{cost}}(\delta) = M(n_1, n_2, k, \delta) - \widetilde{M}(n_1, n_2, k, \omega_d)$. The cost function is then solved for multiple k values using least-squares minimisation for the unknown damping δ [35].

The amplitude and phase angle are identified from the MW integral. The difference between the measurement and the analytically based MW integral is used to set the cost functions for the amplitude and phase angle. The cost function for the amplitude identification is: $F_{\text{cost}}(A) = \left| W(n_1, k, \omega_d, \delta, A) \right| - \left| \widetilde{W}(n_1, k, \omega_d) \right|$, and the cost function for phase-angle identification: $F_{\text{cost}}(\phi) = \phi - \tilde{\phi}(k)$. Both cost functions are solved separately for multiple k values using the least-squares minimization for the unknown amplitude A and phase angle ϕ , respectively [35]. The phase angle $\tilde{\phi}(k)$ is obtained from the argument of the MW integral (5) with:

$$\tilde{\phi}(k) = -\arctan \left(\frac{\Im \left[(-1)^k \widetilde{W}(n_1, k, \omega_d) \right]}{\Re \left[(-1)^k \widetilde{W}(n_1, k, \omega_d) \right]} \right) \quad (12)$$

2.4. Discretisation of light reflectance for the object

This section discusses the basic principle of how the camera sensor converts light reflected from the object into pixel intensities. The aim is to show the main sources of error in the conversion and how the optical flow is registered. For more details on these topics the reader should be referred to [39, 40].

The light reflected from the object (Fig. 2) falls on the photo-sensitive areas of the CMOS sensor, where it is spatially discretised into pixels (the number of photo-sensitive areas on the camera sensor is usually expressed in Mpx). The light at each pixel is converted into an

analogue signal, which is then digitised by the AD converter into unsigned integer values in the range $0 \leq I < 2^Q$, where $Q \in \mathbb{N}$ is the resolution of the AD converter, expressed by the number of bits Q . The process of converting the light reflected from the object into pixel intensity is shown graphically in Fig. 3 [39].

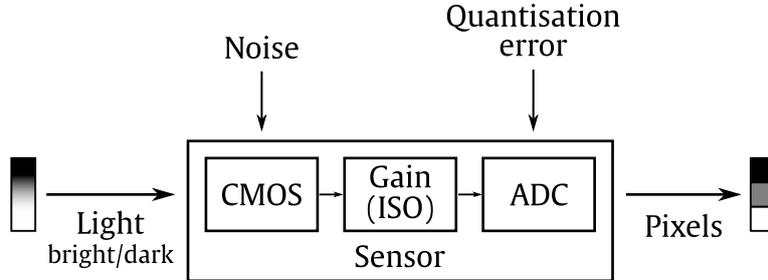


Figure 3: Schematic of the conversion of light into pixel intensities on the sensor of a digital camera with sources of error.

The conversion is not ideal and the signal is contaminated with noise from the CMOS sensor and the AD converter, which is typically 8 or 12 bits in today’s camera sensors. The noise from the CMOS sensor dominates over the quantisation error and comes from various sources, *e.g.*, photoelectron noise, readout noise and dark-current noise [41]. The noise requires a sophisticated model [42]; however, for simplicity it is here represented with Gaussian noise.

The pixels on the camera sensor are usually rectangular and are defined geometrically. Theoretically, the total light intensity measured by a pixel is related to the number of photons captured by that pixel during the exposure time. It is important that the exposure time is as short as possible to avoid blurring effects [28]. A short exposure time can be achieved directly by exposing the object to strong light sources so that enough photons reach the sensor, or by increasing the gain (higher ISO numbers). The first option is more desirable because increasing the gain also increases the noise coming from the sensor, resulting in a lower signal-to-noise ratio.

3. Optical-flow-based EMA

In this section the principles for performing a modal identification directly from pixel intensities are presented, taking into account the influence of the spatial gradient. The registration of the motion from the object $q(x, t)$ to the light intensities $I(x, t)$ is depicted in Fig. 4 for the ideal example. To apply the Morlet-wave integral directly to the pixel intensities, the following assumptions are required: the motion must be within the sub-pixel domain $A \leq 1$ px and there must be a linear relationship between the light intensity and the displacement. It should be noted that the pixel intensities are digitised light intensities, which except in Sec. 4, are treated as continuous.

To identify the modal parameters, a free response is required. When the response of the object $q(x, t)$ is obtained by hitting the modal hammer, the amplitudes at the beginning of

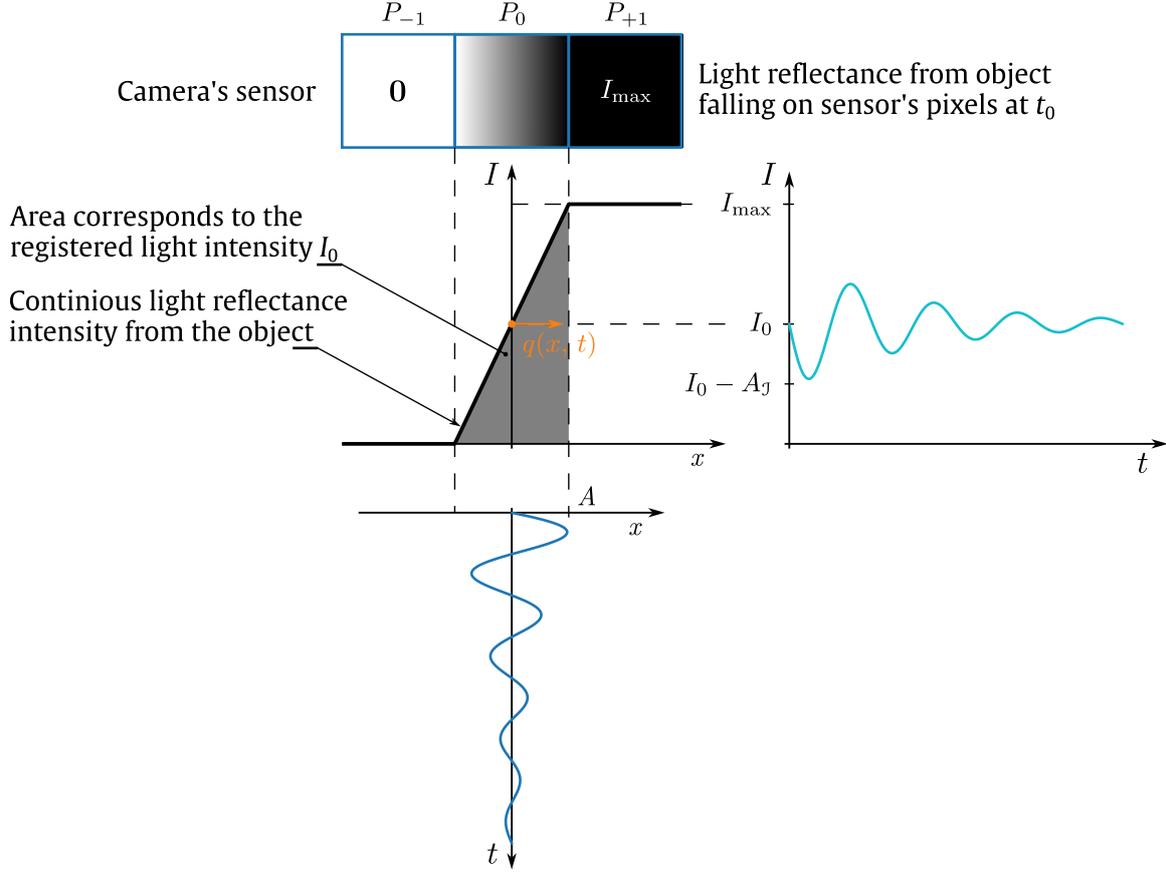


Figure 4: Conversion of the light reflectance from the object into pixel intensities.

the signal may exceed the 1 px assumption. Mitigation involves shifting the integration in Eq. (5) by t_0 :

$$\widetilde{W}(n, k, \omega) = \int_{t_0}^{t_0+T} q(x, t) \psi^*(n, k, \omega, t) dt. \quad (13)$$

Then Eq. (3) is inserted into Eq. (13), obtaining the following expression:

$$\widetilde{W}(n, k, \omega) = \frac{1}{I'_0(x)} \left[I_0(x) \underbrace{\int_{t_0}^{t_0+T} \psi^*(n, k, \omega, t) dt}_{\mathcal{L}} - \int_{t_0}^{t_0+T} I(x, t) \psi^*(n, k, \omega, t) dt \right] \quad (14)$$

where the reference image I_0 is time invariant for the subpixel motion and goes out of the integral, which is now applied to the conjugated Morlet-wave function ψ^* . This leads to the admissibility condition, which is approximately equal to zero (6): $\mathcal{L} \approx 0$. Finally, the following expression is obtained:

$$\widetilde{W}(n, k, \omega) = -\frac{1}{I'_0(x)} \underbrace{\int_{t_0}^{t_0+T} I(x, t) \psi^*(n, k, \omega, t) dt}_{\widetilde{W}_j(n, k, \omega)} \quad (15)$$

In Eq. (15) we can observe that the Morlet wavelet can be applied directly to the pixel intensities to obtain the MW integral $\widetilde{W}_j(n, k, \omega)$. The spatial gradient $I'_0(x)$ can be used to scale them to pixel units:

$$\widetilde{W}(n, k, \omega) = -\frac{1}{I'_0(x)} \widetilde{W}_j(n, k, \omega) \quad (16)$$

The response of the structures contains several modes, and due to the linear property of the Morlet wave (7), each mode can be analysed separately [43, 44].

3.1. The exact identification of the natural frequency

The exact natural frequency is identified from the MW integral (5) using the numerical minimisation defined by Eq. (10) [35]. By inserting Eq. (16) into Eq. (10), the operator $\partial/\partial\omega$ acts directly on the MW integral obtained from the pixel intensities:

$$-\frac{1}{I'_0(x)} \frac{\partial}{\partial\omega} \left| \widetilde{W}_j(n, k, \omega) \right| = 0 \quad (17)$$

It is important to obtain the exact natural frequency on the MW integral and the same parameters n, k as will be used later in the modal identification step [35].

3.2. Identification of the damping ratio

By inserting Eq. (16) into Eq. (11), the spatial gradient is cancelled and the identification of the damping ratio is made directly from the pixel intensities:

$$\widetilde{M}(n_1, n_2, k, \omega_d) = \frac{\left| \widetilde{W}_j(n_1, k, \omega_d) \right|}{\left| \widetilde{W}_j(n_2, k, \omega_d) \right|} \quad (18)$$

using the procedure described in Sec. 2.3.

3.3. Identification of the amplitude

Amplitude is identified from the pixel intensities (16) using the cost function:

$$F_{\text{cost}}(A_j) = \left| W_j(n_1, k, \omega_d, \delta, A_j) \right| - \left| \widetilde{W}_j(n_1, k, \omega_d) \right| \quad (19)$$

where W_j is the analytical MW integral (9) based on the amplitude A_j , which is the amplitude of the harmonic part of the light intensities (4). An absolute value of the spatial gradient is required for the scaling to pixel units:

$$A(x) = \frac{A_j(x)}{|I'_0(x)|} \quad (20)$$

which can be further scaled to engineering units.

3.4. Identification of the phase angle

To correctly identify the phase angle from the pixel intensities, the sign has to be corrected, leading to the following expression for the identification of the phase angle:

$$\tilde{\phi}(k) = -\arctan\left(\frac{\Im\left[-\text{sgn}(I'_0) (-1)^k \widetilde{W}_j(n_1, k, \omega_d)\right]}{\Re\left[-\text{sgn}(I'_0) (-1)^k \widetilde{W}_j(n_1, k, \omega_d)\right]}\right) \quad (21)$$

where two sign corrections can be observed: $-\text{sgn}(I'_0)$. The minus sign comes from Eq. (16), the physical aspect of which can be seen in Fig. 4, where it is clear that in the case of a positive gradient, the displacement in the positive direction at P_0 will be registered as a reduction in the light's intensity. On the other hand, if the spatial gradient would be negative then the same motion would be registered as an increase of light intensities and the sign has to be corrected from the gradient with the $\text{sgn}(I'_0)$. The identification is carried out in the same way as described in Sec. 2.3. Using the identified amplitude and phase angle, residues can be obtained for the i th mode with:

$$R_i(x) = A_i(x) e^{i\phi_i(x)} \quad (22)$$

4. Synthetic experiment

The identification of the modal parameters based on the pixel intensities is demonstrated using the synthesised numerical data. The sensitivity of the modal parameters is tested in terms of light intensity and gradient (Sec. 4.1), and sensor noise and quantisation (Sec. 4.2).

4.1. Light intensity and gradient influence

Here, the light intensity of individual pixels is examined in relation to the spatial gradient. The change in light intensity is caused by the harmonic motion, as shown in Fig. 4. Here, the associated motion—change-in-intensity is examined using a line image with three pixels, as shown in Fig. 5 (together with a corresponding gradient image I'_0). In Fig. 5, the change

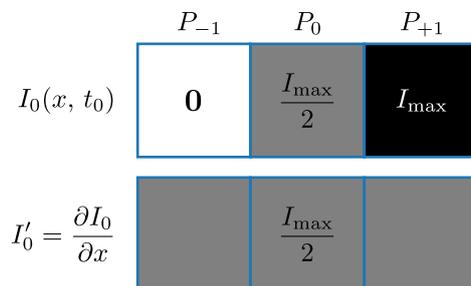


Figure 5: Line image and the corresponding gradient image used for the synthetic experiment at t_0 .

in light intensity at pixel P_0 is based on Eq. (3) and the following expression is obtained:

$$I(x_0, t) = I_0(x_0) - I'_0(x_0) q(t) \quad (23)$$

where $q(t)$ is the free response of the SDOF system, similar to Eq. (4), except that here the amplitude A corresponds to the maximum amplitude of the object, expressed in pixel units. The reference light intensity I_0 is set according to Fig. 5 for the centre pixel P_0 ; in this example for the centre pixel the reference light intensity and the spatial gradient are equal: $I_0 = I'_0 = I_{\max}/2$. The maximum intensity level I_{\max} is defined as a real relative number $0 \leq I_{\max} \leq 1$, where 1 represents the reflected light intensity at maximum illumination.

Gaussian noise is added to the simulated light intensity:

$$I_\sigma(x_0, t) = I(x_0, t) + I_n(\sigma) \quad (24)$$

where $I_n(\sigma)$ is Gaussian noise with a mean of zero and the standard deviation σ . As an example, $\sigma = 13 \times 10^{-4}$ would correspond to a noise of 1-bit on an 8-bit sensor², see Fig. 6.

The sensitivity to light intensity and gradient influence is researched with harmonic motion at subpixel amplitude, *i.e.* $A = [0.1, 0.05, 0.02, 0.01, 0.007, 0.005]$ px. Additionally, for two light-intensity levels at pixel P_0 : $I_0 = [0.5, 0.25]$ are researched. The remaining parameters of the system are: damped natural frequency $\omega_d = 100 \cdot 2\pi \text{ s}^{-1}$, damping ratio $\delta = 0.5\%$ and phase angle $\phi = 1.3$. The camera intensity is sampled at 5000 Hz, the length of the signal is set to $T = 0.6 \text{ s}$. The sampling frequency is set the 50 times above the natural frequency according to the recommendation of Slavič and Boltežar [37]. The simulated response of the light intensity at $A = 0.05 \text{ px}$ against the noise is shown in Fig. 6.

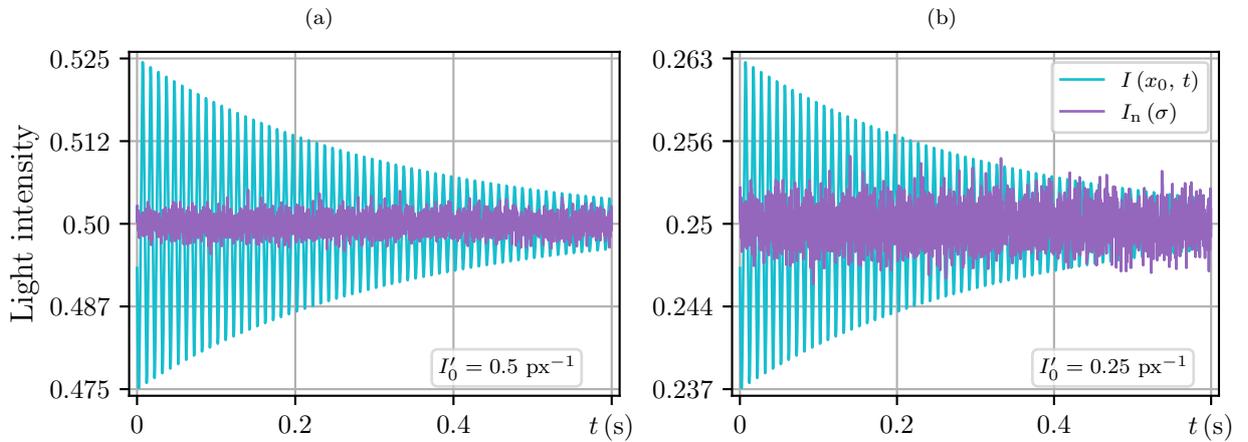


Figure 6: Light-intensity response of the SDOF system with amplitude $A = 0.05 \text{ px}$ against noise, for the light intensities at pixel P_0 : a) $I_0 = 0.5$ and b) $I_0 = 0.25$.

Modal identification is performed with the Morlet-wave (MW) modal package [45] using the default parameter set [35] ($n_1 = 5$, $n_2 = 10$, $k_{l_0} = 10$, $N_k = 10$, $\delta_e = 0.25\%$) on simulated pixel intensities with 15 different realisations of noise. The MW modal method is applied

²The noise, expressed in the number of Q_{noise} bits, is consistent with the resolution of the quantiser Q over three standard deviations of Gaussian noise with: $3\sigma = 2^{Q_{\text{noise}}-1}-Q$.

to the pixel intensities subtracted from the reference image $I_0(x) - I_\sigma(x_0, t)$. The results of the identification are expressed as relative deviations from the theoretical values in Fig. 7, with error bars where the centre represents the mean and the lines indicate the standard deviation from the mean.

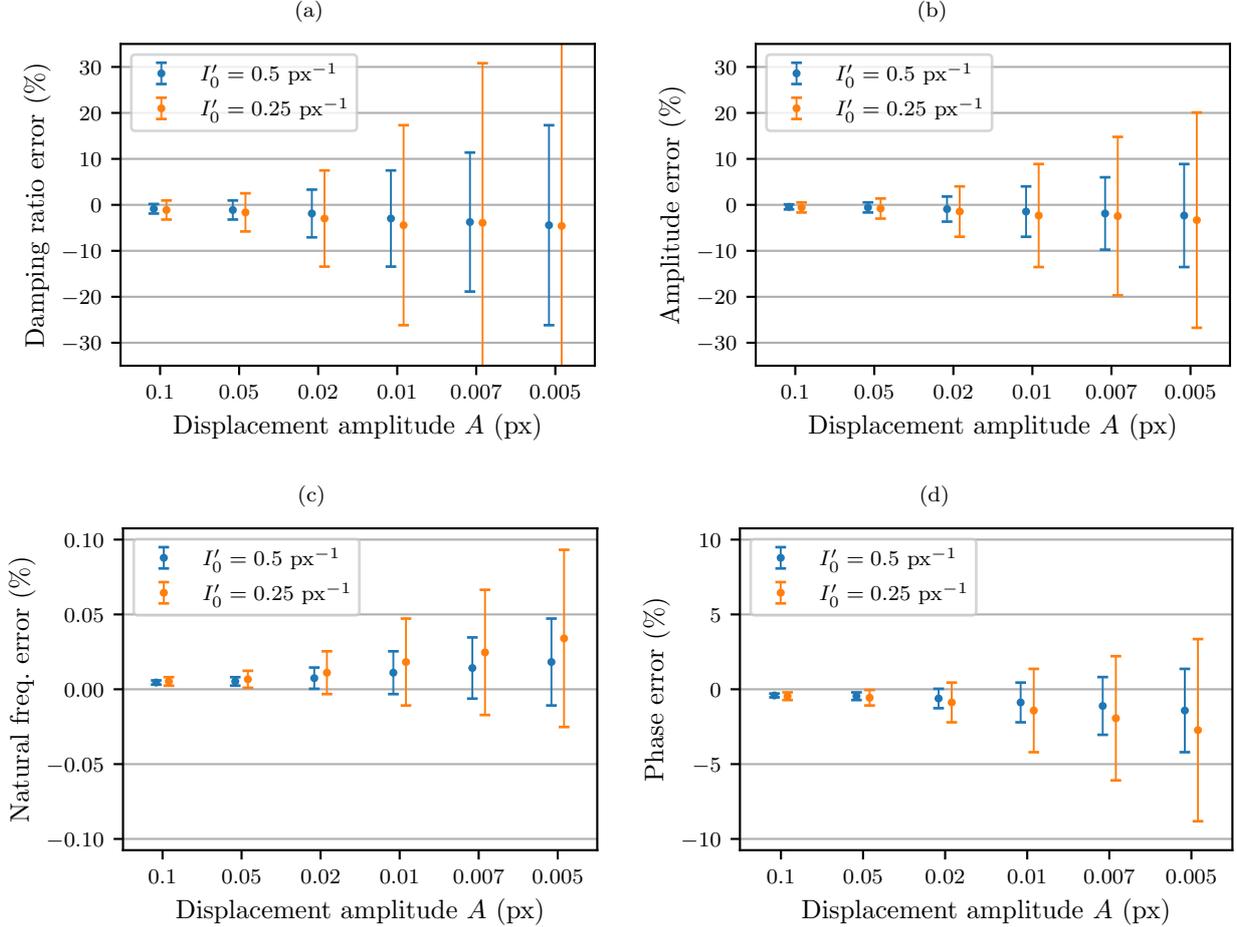


Figure 7: Comparative results of modal identification between two gradient levels on 15 different realisations of the noise for: a) damping ratio, b) amplitude, c) natural frequency and d) phase angle.

4.2. Influence of noise and quantisation on modal identification

The modal identification is tested against sensor noise and quantisation with the simulated gradient case $I'_0 = 0.25 \text{ px}^{-1}$ from Sec. 4.1. Here, the Gaussian noise with two standard deviations $\sigma = 13 \times 10^{-4}$ and $\sigma = 6.5 \times 10^{-4}$ is investigated. To obtain pixel intensities similar to those of a real camera, the light intensity with added noise $I_\sigma(x_0, t)$ (24) is quantised using the following expression:

$$\tilde{I}_\sigma(x_0, t) = \lfloor I_\sigma(x_0, t) \cdot (2^Q - 1) + 0.5 \rfloor \quad (25)$$

for the resolutions $Q = 8$ bit and $Q = 12$ bit (see Fig. 3), where $\lfloor \cdot \rfloor$ is the floor operation. The simulated pixel intensity at amplitude $A = 0.05$ px and for 8/12-bit resolution is shown in Fig. 8. Note: In this section only, pixel intensity is denoted by the tilde $\tilde{I}_\sigma(x_0, t)$ to distinguish it from light intensity $I_\sigma(x_0, t)$.

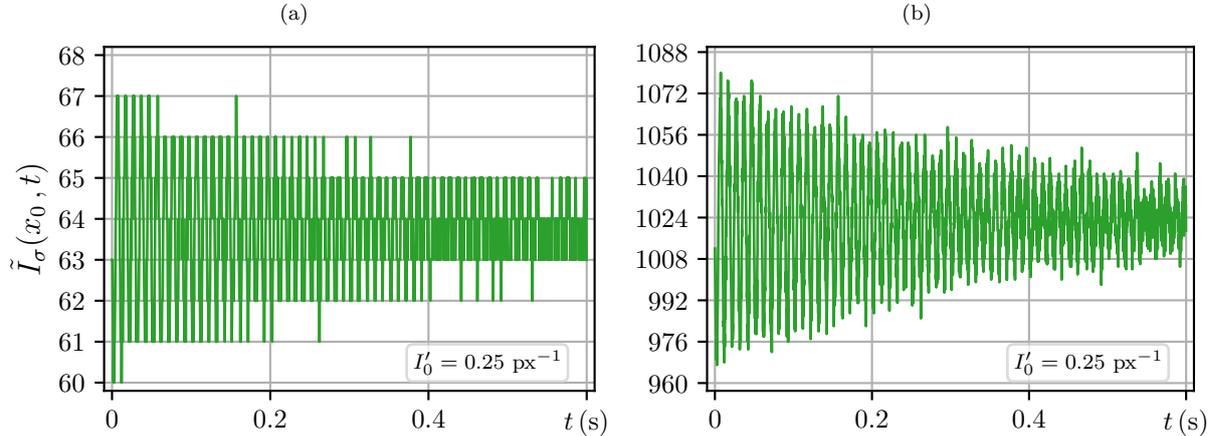


Figure 8: Pixel-intensity response of the SDOF system simulated with amplitude $A = 0.05$ px for: a) $Q = 8$ -bit and b) $Q = 12$ -bit quantisation resolution.

The MW modal method [45] is used with the same parameters as in Sec. 4.2. The simulated response $\tilde{I}_\sigma(x_0, t)$ is subtracted from the reference intensities \tilde{I}_0 after quantisation, which is obtained as a time average of $\tilde{I}_\sigma(x_0, t)$. The results are expressed as relative deviations from the theoretical values in Fig. 9, with error bars where the centre represents the mean and the lines indicate the standard deviation from the mean.

4.3. Discussion

The results from Sec. 4.1 in Fig. 7 show the influence of the gradient on the identification of the modal parameters in relation to the sensor noise. The example of the line image (Fig. 5) is an ideally registered, sharp edge for two illumination cases. The desired effects are demonstrated using the response of SDOF system as an example, and due to the linearity of the Morlet wave integral (7), the same effects are expected for the MDOF systems. The first case represents the ideally illuminated edge where the light intensities through three pixels range linearly from 0 to 1. The second case is half of the ideal, which is closer to the real illumination. Such a distribution of pixel intensities is required to determine the gradient, unlike the continuous case (see Fig. 3). Therefore, the gradient directly affects the sensitivity of the registered movement due to its scaling effect (23), and it is desirable to have a higher gradient. On the other hand, the sensor noise depends on the light intensities (the higher the light, the higher the noise), which is also camera dependent [46]. For simplicity, it is assumed that the sensor emits the same amount of noise and two discrete noise levels ($\sigma = 13 \times 10^{-4}$ and $\sigma = 6.5 \times 10^{-4}$) were used instead. The 15 different realisations of the noise are chosen to replicate the number of pixels from the subset as in Sec. 5.2. For the

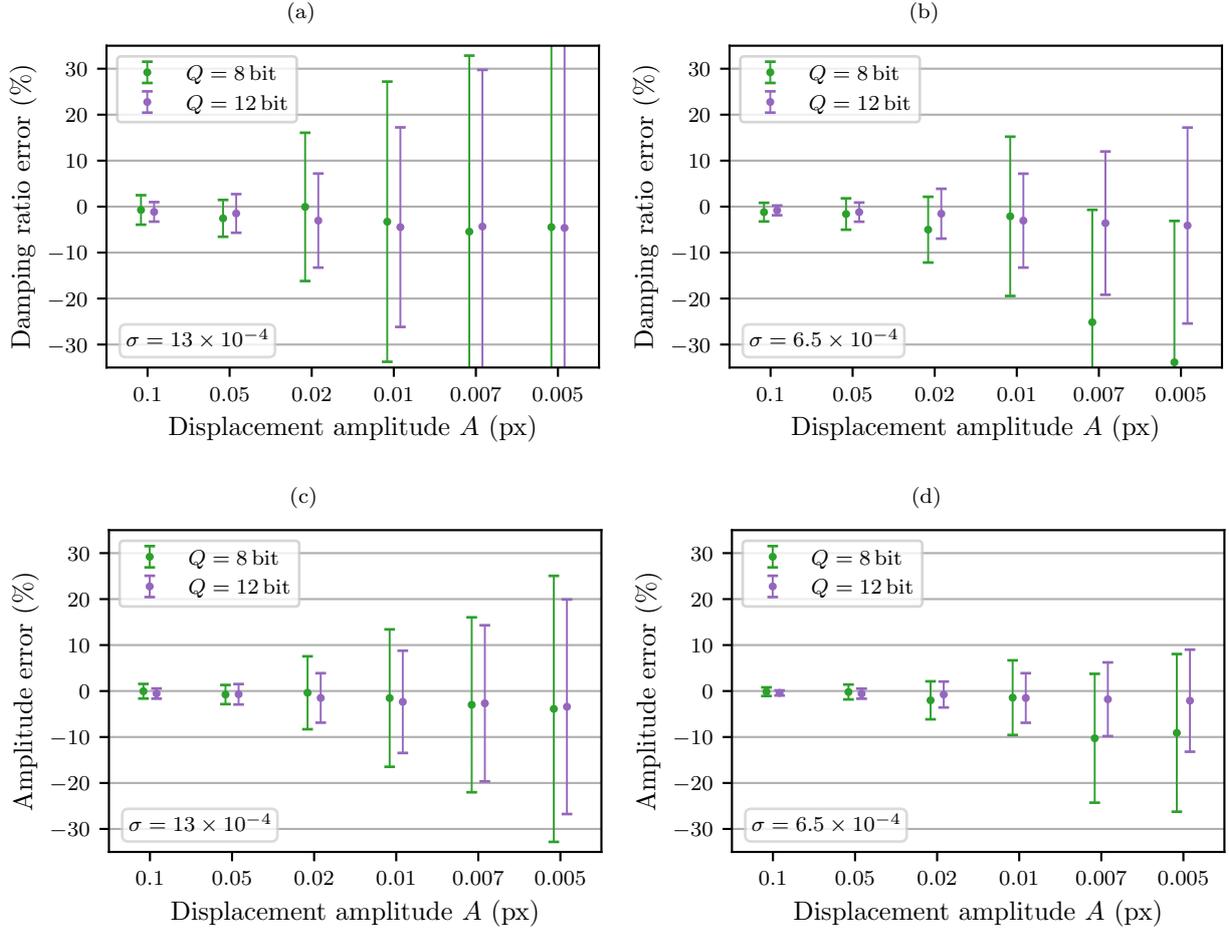


Figure 9: Results of the damping ratio and the amplitude identification between two quantisation resolutions on 15 different realisations of the noise for standard deviation of noise: a) and c) $\sigma = 13 \times 10^{-4}$, b) and d) $\sigma = 6.5 \times 10^{-4}$.

applied noise level (1-bit noise with an 8-bit sensor), all the modal parameters are identified accurately according to the mean values, with an error up to 5% for the damping and up to 3% for the phase. However, the standard deviation of the damping is up to $\pm 10\%$ from the mean for amplitudes greater than 0.02 px, indicating that the damping from a single pixel can be identified with an error of at most 10%. In contrast, at smaller amplitudes (and with the noise level kept the same as before), the standard deviation from the mean increases from $\pm 20\%$ at $A = 0.01$ px to $\pm 45\%$ at $A = 0.005$ px, indicating that results from more pixels are required to increase the accuracy of the identification. The mean-error value of the amplitude is similar to that of the damping, and is below $\pm 5\%$ for all the amplitudes tested, and the standard deviation from the mean has the same trend as the damping, only with the errors halved.

The results from Sec. 4.2 in Figs. 9a and 9c show no significant change in the mean error due to quantisation when compared to the results from Sec. 4.1 (Figs. 7a and 7b) for the

same gradient case ($I'_0 = 0.25 \text{ px}^{-1}$). The damping and amplitude are identified with similar average errors for both resolutions (up to 5% at amplitude $A = 0.005 \text{ px}$). The deviation of the results for the quantisation with $Q = 12$ -bit resolution did not change significantly (below $\pm 1\%$) compared to the results in Figs. 7a and 7b (see orange error bars). For $Q = 8$ -bit quantisation, the deviation of the results increased from 6% at $A = 0.02 \text{ px}$ to 11% at $A = 0.005 \text{ px}$, for higher amplitude cases ($A = [0.1, 0.05] \text{ px}$) the deviation was below 3%. When the standard deviation of the noise is halved ($\sigma = 6.5 \times 10^{-4}$), the deviation of the results halves in all cases, but the mean error of the damping and amplitude for the 8-bit quantisation increases at amplitudes $A < 0.01 \text{ px}$, which can be observed in Figs. 9b and 9d (see green error bars).

Quantisation with 8-bit resolution can register the minimum light-intensity amplitude $A_{j,\min} \approx 0.004$. Scaling with the gradient $I'_0 = 0.25 \text{ px}^{-1}$ using Eq. (20), the minimum amplitude $A_{\min} = 0.016 \text{ px}$ is obtained. Therefore, the amplitude $A = 0.02 \text{ px}$ for the gradient $I'_0 = 0.25 \text{ px}^{-1}$ is a limiting case for the 8-bit quantisation shown in the case with noise std. dev. $\sigma = 6.5 \times 10^{-4}$ (Figs. 9b and 9d). In contrast, the higher noise ($\sigma = 13 \times 10^{-4}$) improved the averaged results for the cases below A_{\min} (Figs. 9a and 9c). The results show that quantisation has no significant effect on the 12-bit resolution compared to the errors due to noise. The higher noise level also corresponds to 5-bit of noise on 12-bit resolution, which is to be expected in the experimental example. To achieve the same quantisation effect at 12-bit resolution (Figs. 9b and 9d), selected subpixel amplitudes A should be proportionally lowered (for gradient intensity $I'_0 = 0.25 \text{ px}^{-1}$), while maintaining the same signal-to-noise ratio, the noise should be decreased to $\sigma = 0.41 \times 10^{-4}$. The results of the quantisation of natural frequency and phase were omitted because the quantisation errors did not have a significant impact compared to the results in Sec. 4.1 (Figs. 7c and 7d).

The results were obtained with the sampling rate recommended by Slavič and Boltežar [37], which are preferred for high noise and a low dynamic range data. Sampling rates such as $5\times - 10\times$ (of the analysed frequency ω) can be used for harmonic amplitudes $A > 0.02 \text{ px}$ at 8-bit (gradient intensity $I'_0 = 0.25 \text{ px}^{-1}$), which corresponds to $A > 0.001 \text{ px}$ at 12-bit.

In addition, this study found that approximating the admissibility condition (6) in Eq. (14) causes errors in the identified damping ratio and phase angle for values $n < 7$ and $k < 40$. The error is mitigated by subtracting the referenced light intensity I_0 .

5. Experimental research

In this section, modal identification is performed experimentally on the basis of a single pixel. First, the identification of the modal parameters based on a single pixel in terms of gradient intensities is demonstrated, including a comparison with the results obtained from displacements identified with a digital image correlation (DIC) method. Then the modal identification of the whole structure is performed based on the information of the single pixel and DIC method is used to compare mode shapes.

The laboratory, freely supported, steel beam ($w \times h \times d = 500 \times 15 \times 30 \text{ mm}$) was used to perform the experiment. The beam has a notch ($w \times h \times d \approx 2 \times 8 \times 30 \text{ mm}$)

in the middle on the underside (see Fig. 10). Excitation involved a single hit with a modal hammer (KISTLER 9726A20000) and the response was recorded simultaneously with the accelerometer (PCB J320C03) and the high-speed camera (FASTCAM SA-Z type 2100K-M-64GB, 12 bit greyscale sensor). The front of the beam was covered with a black-and-white stripe recorded at 100 kFPS for a duration of 1 s with the electronic shutter set to 1/119149 s. The excitation force and the response of the acceleration signals were sampled with a NI-9234 data-acquisition (DAQ) device at 51.2 kHz. The camera and the DAQ device were synchronised with an electrical pulse generated when the beam was hit with a modal hammer, similar to [47], except that in this research a single laptop was used to run the camera and DAQ software. The experimental setup is shown in Fig. 10, and the signal of the hammer impact is shown in Fig. 11.

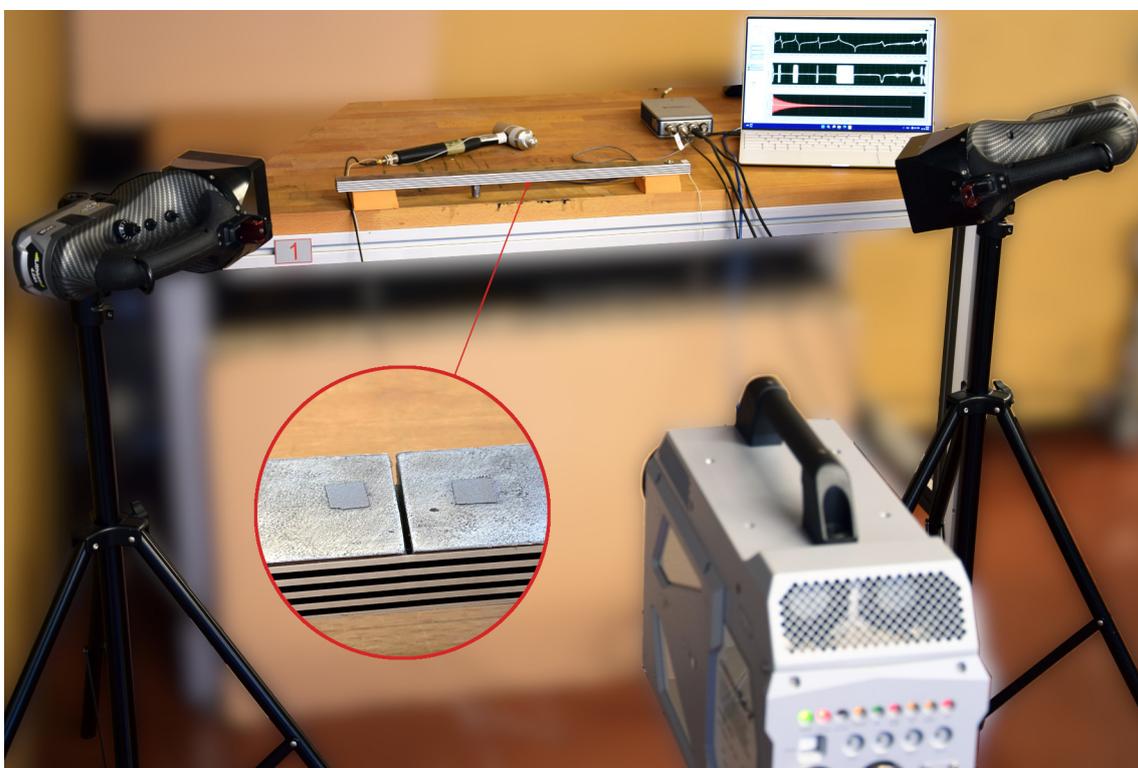


Figure 10: Experimental setup. Detail in red circle shows the notch on the underside of the beam.

The modal identification was performed directly on the pixel intensities in a time frame of $T_f = 0.5$ s, which was shifted for $t_0 = 75$ ms to avoid rigid-body and low-frequency oscillations of more than 1 px at the beginning of the recording caused by relatively soft supports (polyurethane foam blocks). Fig. 12 shows the first image taken by the high-speed camera, which also shows the positions of the hammer impact and the accelerometer. The pixels used to demonstrate the modal identification using a single pixel are marked with a cyan line, and for modal identification of the whole model, the pixels with the orange rectangle of size 30×994 px are selected, both shown in Fig. 12.

The reference image $I_0(x, y)$ was obtained by averaging all the pixels within the orange

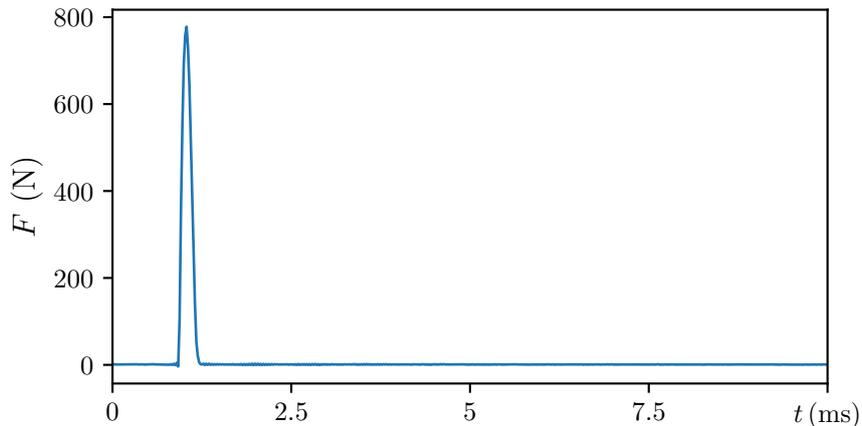


Figure 11: Impact of the modal hammer.

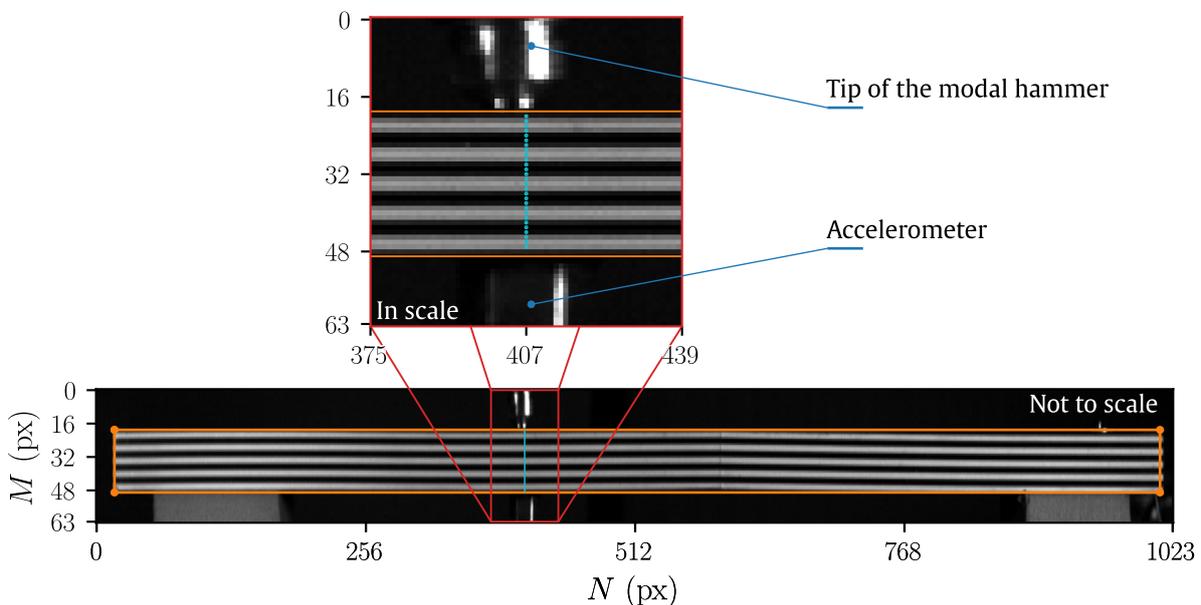


Figure 12: Full frame, taken with a high-speed camera. The cyan line in the full-frame and the cyan dots in the detail image mark the pixels used for the single-pixel modal identification, the orange rectangle (30×994) px marks the range of pixels used for modal identification of the whole model.

rectangle (Fig. 12) in the defined time frame T_f . The averaging yields a single image that is less affected by noise due to the harmonic response in the subpixel range, which is then used to obtain the gradient intensities (Sec. 2.1). Next, the spatial gradient image in the direction of motion was obtained: $I'_0 = \partial I_0(x, y)/\partial y$, using the second order accurate central differences. The pixel intensities and the corresponding spatial gradient of the pixels located at pixel coordinate $N = 407$ in the reference image (cyan line in Fig. 12) are shown in the bar plot in Fig. 13. The pixels at the edge of the beam have been omitted due to errors.

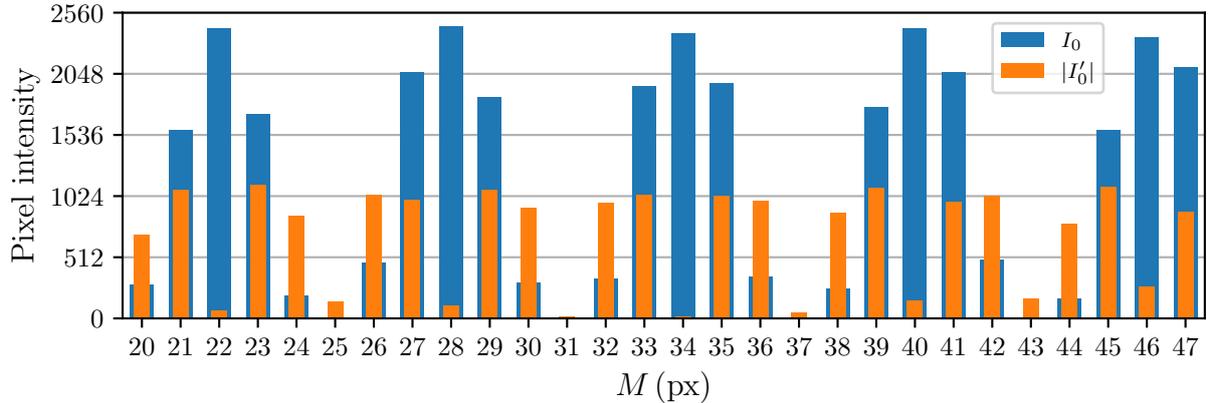


Figure 13: Pixel intensities compared to the spatial gradient from the reference image I_0 (30×994) at the selected column ($N = 407$).

5.1. Single-pixel modal identification

Modal identification based on a single pixel is demonstrated here using the example of 28 pixels (detail frame in Fig. 12 – cyan dots) whose intensities and spatial gradient are shown in Fig. 13. The response of the pixel intensity both in the time domain $I(t)$ and in the frequency domain $\hat{I}(\omega)$ is shown in Fig. 14 using the example of the pixel at coordinates $M = 39$, $N = 407$.

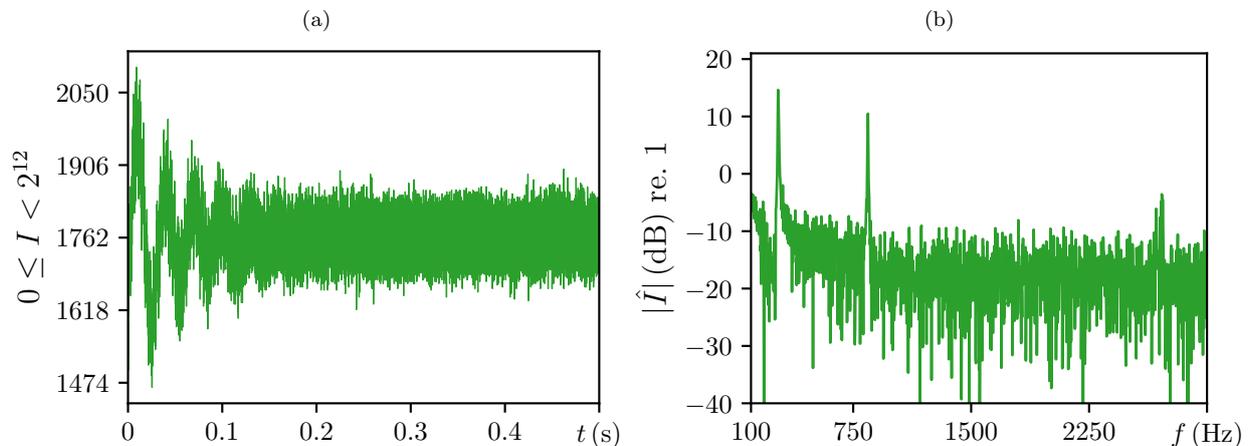


Figure 14: Response of the selected pixel in the (a) time and (b) frequency domain.

Displacements are identified using the DIC method as implemented in the Python package pyIDI [48]. The region-of-interest (ROI) is centred on (3×9) px at 10 locations horizontally centred on pixel coordinate $N = 407$ and vertically centred on pixel coordinates $M = [21, 23, 27, 29, 33, 35, 39, 41, 45, 47]$, to include pixels with high gradient intensities in the vertical direction (3 px) and in the horizontal direction with 9 px, which was minimally required to identify displacements at all selected points.

Modal identification was performed for the second mode. The initial natural frequency is selected from the magnitude plot as Fig. 14b which is $\hat{f}_2 = 844$ Hz and the MW modal method [45] was applied to pixels, displacements identified with DIC and the accelerometer responses using the following parameters: $n_1 = 5$, $n_2 = 10$, $k_{lo} = 40$ and $\delta_e = 0.2\%$. Selected pixel intensities were subtracted from the reference image to mitigate the influence of the approximation in Eq. (14) as in Sec. 4 and the sign is corrected with gradient $I'_0(x, y)$ required for phase identification in Eq. (21): $I_0(x, y) - I(x, y, t) \cdot \text{sgn}(I'_0(x, y))$. The results of the modal identification are shown in Fig. 15 and are compared to the results of DIC and the accelerometer.

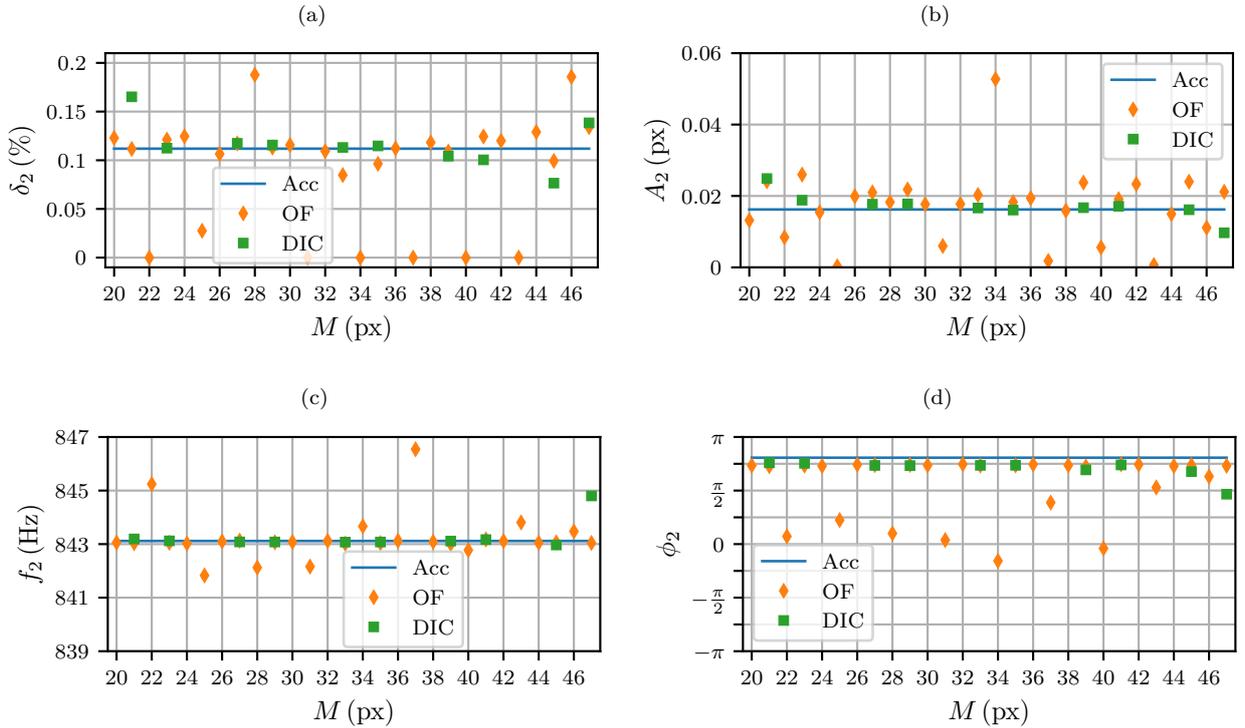


Figure 15: Single-pixel modal identification results compared to results obtained from DIC and accelerometer for: a) damping ratio, b) amplitude, c) natural frequency and d) phase angle.

Pixels selected for modal identification oscillate uniformly due to the small oscillations of the beam and form a subset. The results in Fig. 15 show that the modal parameters obtained from the pixels at the row coordinates $M = [22, 25, 28, \dots, 46]$ are different from the results at the other rows, including the accelerometer results. In Fig. 13, we can see that the spatial gradient at these coordinates is significantly lower than at the others. As explained in Sec. 4, a low spatial gradient means a low sensitivity of the motion registration to light intensities and therefore these pixels can be omitted from the identification.

5.2. Modal identification of the whole model

In order to perform the modal identification along the beam, the reference image I_0 is divided into 994 subsets, where each subset represents 28 pixels in each column N of the reference image $I_0(x, y)$. Unlike in Sec. 5.1, where the identification was performed for all pixels in the subset, here the pixels suitable for modal identification are selected as:

- The best pixel - a single pixel with the highest spatial gradient in the subset is selected
- Average pixel - 14 pixels with the highest gradient in the subset are selected and the results are averaged

Additionally, the displacements from the same pixel responses (Fig. 12) are identified using the DIC method implemented in the python package pyIDI [48]. The identification of displacements is performed on ROI of size (21×21) px, centred between all pixel coordinates in the range $9 \leq N \leq 985$ and $M = 15$, with a total of 977 ROIs.

The MW modal method [45] is applied in the same way as in Sec. 5.1, using the same parameters for all modes. The initial natural frequencies are selected from the power spectral density $P_j(\omega) = \hat{I}_j^*(\omega) \hat{I}_j(\omega) / T$ obtained by averaging the PSDs of the pixels with the highest gradient in all subsets j : $\bar{P}(\omega) = \frac{1}{N} \sum_{j=1}^{N=994} P_j(\omega)$. The averaged PSD $\bar{P}(\omega)$ with selected natural frequencies is shown in Fig. 16.

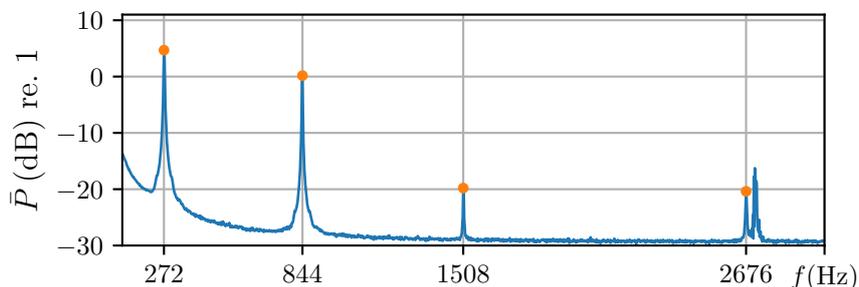


Figure 16: Averaged PSD from all subsets. The orange dots mark the peaks of the natural frequencies. The peak on the far right corresponds most closely to the first torsion mode.

The mode shapes are obtained from the identified amplitude A_i and phase angle ϕ_i using Eq. (22). To simplify graphical representation of the identified mode shapes they are collapsed on the imaginary axis with $r_i = \Im[R_i e^{-i\theta_i}]$, where $\theta_i = [67, 44, -3, -35]$ deg. Mode-shapes are compared for the first and second cases, in Fig. 17 for the first two modes and in Fig. 18 for the remaining modes. In addition to the mode shapes identified by averaging the pixels in Figs. 17b, 17d, 18b, 18d, the mode shapes identified from the DIC data are added for comparison. The identification is done in the same way as for the pixel data. The results of the natural frequencies and damping ratios for the single pixel and the average pixel approaches are presented in Tab. 1. The results are spatially weighted, as described by Tomac and Slavič [47]. A comparison with the identification results of the acceleration signal, which served as a reference value, is also given.

Table 1: Results of identification of natural frequencies and damping ratios for the best pixel and averaged pixel approach, including a comparison with results obtained from the accelerometer data.

Mode:	1 st		2 nd		3 rd		4 th	
	best pixel	averg. pixel						
$f_{d,acc}$ (Hz)	272.9		843.1		1508.9		2676.4	
f_d (Hz)	272.9	272.9	843.1	843.1	1508.7	1508.8	2676.9	2677.2
Error: (%)	0.0	0.0	0.0	0.0	0.01	0.01	-0.02	-0.03
δ_{acc} (%)	0.519		0.112		(0.062)		0.079	
δ (%)	0.519	0.52	0.112	0.112	(0.121)	(0.084)	0.117	0.085
Error: (%)	0.2	-0.1	0.1	0.2	(-113.2)	(-36.6)	-58.7	-7.7

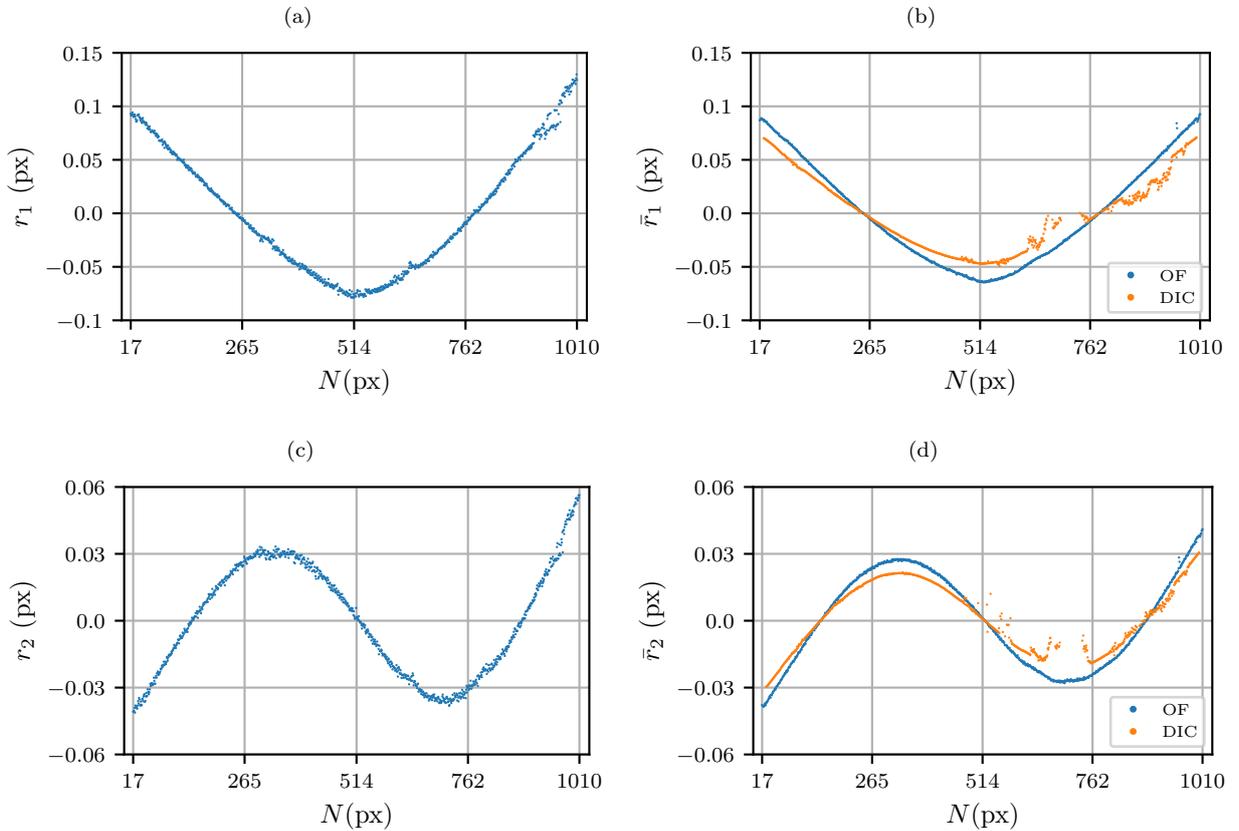


Figure 17: Mode shapes for the first and second natural frequencies based on: (a), (c) the best pixel (OF) and (b), (d) the average pixel (OF) and DIC.

5.3. Discussion

The results of Sec. 5.1, shown graphically in Fig. 15 represent a comparison of the modal identification carried out directly from the pixel intensities, the displacements obtained with

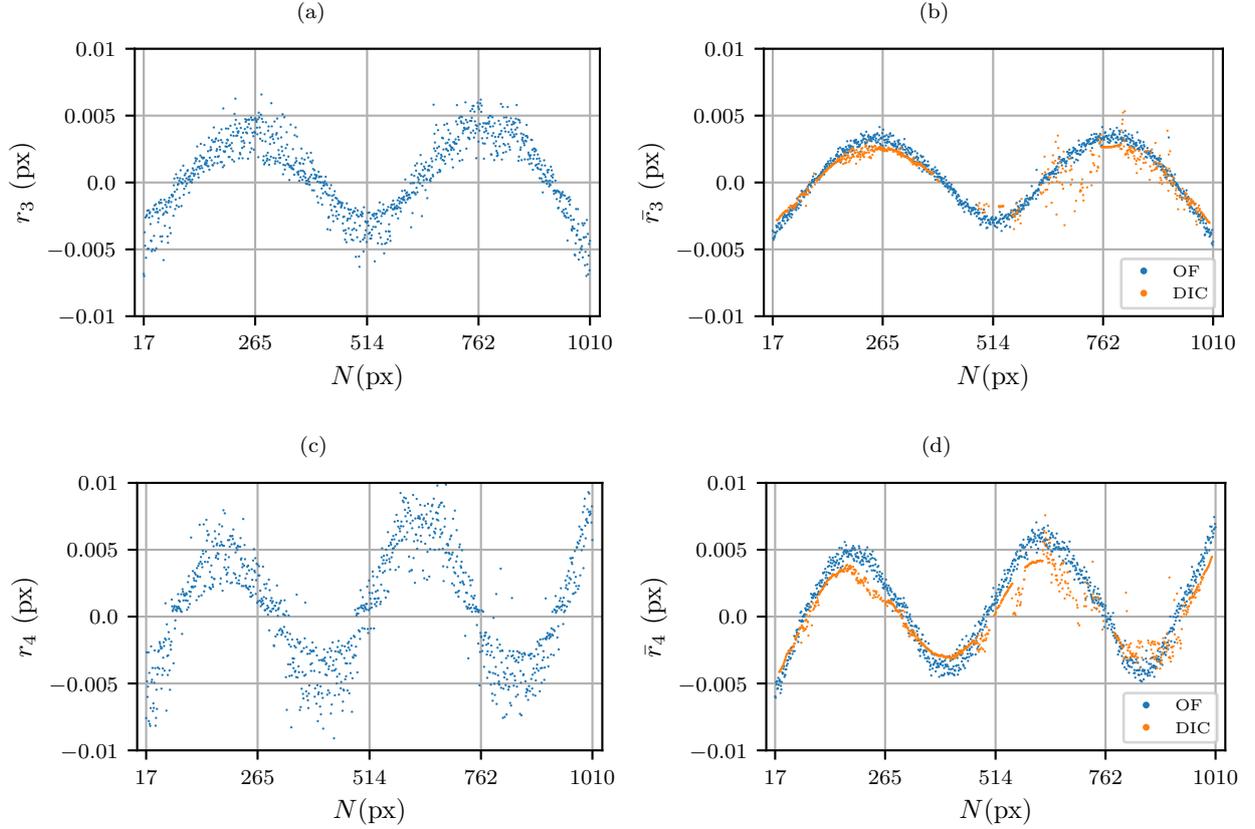


Figure 18: Mode shapes for the third and fourth natural frequencies based on: (a), (c) the best pixel (OF) and (b), (d) the average pixel (OF) and DIC.

DIC and the accelerometer (high-dynamic range sensor) used as a reference. The mean error values from the accelerometer reference are discussed. All pixels that have high gradient intensity level at pixel coordinate $N = 407$ are selected, which corresponds to $I'_0 > 700 \text{ px}^{-1}$ (see. Fig. 13) resulting in total of 19 points. For the DIC the mean values are determined from all 10 locations with ROI (3×9) px. Damping is identified with slightly smaller error from the pixels intensities at 2%, then from the DIC where it is 3.5%. The amplitude is identified with 22% error, while using DIC it is identified with error of 6%. The high amplitude error identified from the pixels is caused by to the accuracy of the calculated gradient, which is determined from the intensities of the neighbouring pixels. The natural frequency is identified with an accuracy of $< 0.01\%$ error and the phase angle is identified with an approximate error of 10%, for both pixels and DIC.

The results of Sec. 5.2 in Tab. 1 show that the natural frequencies can be accurately identified from the response data of a single pixel with a small error (up to 0.02%) and that adding more pixels does not improve the result significantly. The damping, on the other hand, is accurately identified in both cases with an error of $\pm 0.2\%$ for the first and second modes of vibration, and adding more pixels does not significantly improve the result. The

third and fourth modes are at relatively high frequencies and are masked by noise, as can be seen in Fig. 14b, where only the first two modes are clearly visible. However, when the data from more pixels were used, they become distinguishable from the noise, as can be seen in Fig. 16. The same behaviour is observed here in Tab 1, where the damping result of the third mode improved by 68 % (from 113 % to 37 %) and 87 % (from 59 % to 8 %) for the fourth mode when results from more pixels were used. The relatively poor accuracy of the third mode compared to the fourth mode is due to the fact that the third mode was excited near the node, resulting in a higher uncertainty in the accelerometer result.

The identification of mode shapes was similar to that of the damping ratio. In the case of the best pixel, the mode shapes for the first mode (Fig. 17a) and the second mode (Fig. 17c) were clearly identified. The average pixel approach improves the MSE³ of both mode shapes by about 5 %, as can be seen in Figs. 17b and Fig. 17d. The third (Fig. 18a) and fourth (Fig. 18c) mode shapes are more noisy, based on the best-pixel approach. The average-pixel approach improves the results by approx. 25 %, as can be seen in Figs. 18b and 18d. From Figs. 17 and 18 the amplitude of the displacement is approx. ± 0.1 px for the first mode and approx. ± 0.04 px for the second mode, while it is approx. ± 0.004 px for the third mode and approx. ± 0.005 px for the fourth mode. From the magnitude plot of the single-pixel response in Fig. 14 it can be seen that the displacement from the third and fourth modes falls deep into the noise of the camera.

The mode shapes identified from the DIC data agree well with the data from OF on the left side of the beam (Fig. 12), while they are in error on the right side, which can be seen in Figs. 17b, 17d, 18b, 18d. The DIC algorithm has succeeded in identifying the displacements at all selected points. However, they are of lower quality on the right side, most likely due to the pattern used, which is not optimal for DIC and lightening conditions. The amplitude difference observed on the left side, expressed with MSE, is approx. 25 % for the first and second modes and approx. 50 % for the third and fourth. This is due to the accuracy of the gradient calculation from the pixel intensities, as in the Sec. 5.1. The second-order central difference algorithm uses the values of the neighbouring pixels to determine the gradient, which would be fine if the sharp edge was represented with one pixel, as in Fig. 5. However, in Fig. 13 it can be observed that the sharp edges of the pattern used (Fig. 12) are registered across two pixels, caused by the blurring effects of focus and shutter [28], leading to errors in the gradient intensities due to noise. If the gradient intensities $|I'_0|$ were calculated separately (20) by simply subtracting the low intensity value from the high intensity value divided by 2 px, *e.g.* for the pixel intensities in Fig. 13 between pixel coordinates $25 < M < 28$: $|I'_0| = (I_0(28) - I_0(25))/2$, the MSE would be approx. 5 % for the first two modes, 20 % for the third mode and 35 % for the fourth mode.

Numerical processing of the video data with DIC (using the python package pyIDI [48]) and modal identification (using the MorletWaveModal python package [45]) in Sec. 5 was performed on a computer with AMD Ryzen Threadripper 2970WX 24-core processor using parallelisation of ‘for’ loops on 14 cores. Identifying the displacements using the DIC method

³The improvement of the mode shapes is expressed with a mean square error (MSE) [49], averaged mode-shapes are taken for reference.

on ROI (21×21) px at 977 locations in Sec. 5.2 took approx. 2.5 h and another 5 min for modal identification. The time needed to identify the modal parameters for the second case (14 best pixels) in Sec. 5.2 with a total of 13916 px was approx. 1 h.

6. Conclusion

The identification of modal parameters based on pixel intensities from high-speed video recordings using the Morlet-wave modal method is explored. Identification requires a sub-pixel level response of the camera-recorded surface with high gradient features. For mode shapes to be identified in physical units, the gradient image is required. In contrast, the gradient is not required for the identification of the damping ratio and natural frequencies. The sensitivity of the identification based on light intensities is tested in terms of the gradient level under the influence of noise and quantisation. It is shown that by reducing the gradient, the sensitivity decreases and the errors due to noise dominate when using a 12-bit quantisation, while with 8 bit there is an additional influence from the quantisation error for displacements below 0.01 px. In the experimental example it is shown that the response of a single pixel can be used for a modal identification up to 0.04 px, while for displacements down to 0.004 px, the addition of results from more pixels in the selected subset in the modal domain is required to improve identification. The comparison with DIC showed a difference in the identified amplitude with OF, suggesting that more attention needs to be paid to the methodology used to calculate the spatial gradient intensity. With this study it is shown that the Morlet-wave modal method enables modal identification from high-speed video recordings in the same way as it is classically performed with high-dynamic-range sensors.

Acknowledgement

The authors gratefully acknowledge partial financial support from the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie Grant Agreement No. 101027829 and the Slovenian Research Agency (N2-0144).

References

- [1] J. Baqersad, P. Poozesh, C. Niezrecki, P. Avitabile, Photogrammetry and optical methods in structural dynamics – a review, *Mechanical Systems and Signal Processing* 86 (2017) 17–34. URL: <https://doi.org/10.1016/j.ymsp.2016.02.011>. doi:10.1016/j.ymsp.2016.02.011.
- [2] A. Sarrafi, Z. Mao, C. Niezrecki, P. Poozesh, Vibration-based damage detection in wind turbine blades using phase-based motion estimation and motion magnification, *Journal of Sound and Vibration* 421 (2018) 300–318. doi:10.1016/j.jsv.2018.01.050.
- [3] S. Cao, H. Nian, J. Yan, Z. Lu, C. Xu, Modal analysis and damage localization in plate-type structures via tdd and pe methods based on the data of an integrated highspeed camera system, *Mechanical Systems and Signal Processing* 178 (2022) 109309. doi:10.1016/j.ymsp.2022.109309.
- [4] K. Zaletelj, J. Slavič, M. Boltežar, Full-field dic-based model updating for localized parameter identification, *Mechanical Systems and Signal Processing* 164 (2022) 108287. URL: <https://doi.org/10.1016/j.ymsp.2021.108287>. doi:10.1016/j.ymsp.2021.108287.
- [5] P. Gardonio, R. Rinaldo, L. D. Bo, R. D. Sal, E. Turco, A. Fusiello, Free-field sound radiation measurement with multiple synchronous cameras, *Measurement* 188 (2022) 110605. URL: <https://doi.org/10.1016/j.measurement.2021.110605>. doi:10.1016/j.measurement.2021.110605.
- [6] P. Gardonio, G. Guernieri, E. Turco, L. D. Bo, R. Rinaldo, A. Fusiello, Reconstruction of the sound radiation field from flexural vibration measurements with multiple cameras, *Mechanical Systems and Signal Processing* 195 (2023) 110289. doi:10.1016/J.YMSSP.2023.110289.
- [7] T. Willems, F. S. Egner, Y. Wang, M. Kirchner, W. Desmet, F. Naets, Time-domain model identification of structural dynamics from spatially dense 3d vision-based measurements, *Mechanical Systems and Signal Processing* 182 (2023) 109553. URL: <https://doi.org/10.1016/j.ymsp.2022.109553>. doi:10.1016/j.ymsp.2022.109553.
- [8] B. Pan, Digital image correlation for surface deformation measurement: historical developments, recent advances and future goals, *Measurement Science and Technology* 29 (2018) 082001. doi:10.1088/1361-6501/aac55b.
- [9] H. Schreier, J.-J. Orteu, M. A. Sutton, *Image Correlation for Shape, Motion and Deformation Measurements*, Springer US, 2009. doi:10.1007/978-0-387-78747-3.
- [10] A. Zanarini, Full field optical measurements in experimental modal analysis and model updating, *Journal of Sound and Vibration* 442 (2019) 817–842. doi:10.1016/J.JSV.2018.09.048.
- [11] Ángel J. Molina-Viedma, L. Felipe-Sesé, E. López-Alba, F. A. Díaz, Comparative of conventional and alternative digital image correlation techniques for 3d modal characterisation, *Measurement* 151 (2020) 107101. URL: <https://doi.org/10.1016/j.measurement.2019.107101>. doi:10.1016/J.MEASUREMENT.2019.107101.
- [12] J. Javh, J. Slavič, M. Boltežar, The subpixel resolution of optical-flow-based modal analysis, *Mechanical Systems and Signal Processing* 88 (2017) 89–99. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0888327016304770>. doi:10.1016/j.ymsp.2016.11.009.
- [13] R. Huňady, M. Hagara, A new procedure of modal parameter estimation for high-speed digital image correlation, *Mechanical Systems and Signal Processing* 93 (2017) 66–79. URL: <https://doi.org/10.1016/j.ymsp.2017.02.010>. doi:10.1016/j.ymsp.2017.02.010.
- [14] M. H. Richardson, D. L. Formenti, Parameter estimation from frequency response measurements using rational fraction polynomials, volume 1, 1982, pp. 167–186.
- [15] R. Allemang, D. Brown, A unified matrix polynomial approach to modal identification, *Journal of Sound and Vibration* 211 (1998) 301–322. doi:10.1006/jsvi.1997.1321.
- [16] Y. Yang, C. Dorn, T. Mancini, Z. Talken, G. Kenyon, C. Farrar, D. Mascareñas, Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification, *Mechanical Systems and Signal Processing* 85 (2017) 567–590. doi:10.1016/j.ymsp.2016.08.041.
- [17] N. Wadhwa, M. Rubinstein, F. Durand, W. T. Freeman, Phase-based video motion processing, *ACM Trans. Graph.* 32 (2013). URL: <https://doi.org/10.1145/2461912.2461966>. doi:10.1145/2461912.2461966.

- [18] A. Maćkiewicz, W. Ratajczak, Principal components analysis (pca), *Computers & Geosciences* 19 (1993) 303–342. doi:10.1016/0098-3004(93)90090-R.
- [19] J. V. Stone, Blind source separation using temporal predictability, *Neural Computation* 13 (2001) 1559–1574. URL: <https://doi.org/10.1162/089976601750265009>. doi:10.1162/089976601750265009.
- [20] M. Feldman, Hilbert transform in vibration analysis, *Mechanical Systems and Signal Processing* 25 (2011) 735–802. doi:10.1016/J.YMSSP.2010.07.018.
- [21] J. Javh, J. Slavič, M. Boltežar, High frequency modal identification on noisy high-speed camera data, *Mechanical Systems and Signal Processing* 98 (2018) 344–351. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0888327017302637>. doi:10.1016/j.ymsp.2017.05.008.
- [22] H. V. D. Auweraer, W. Leurs, P. Mas, L. Hermans, Modal parameter estimation from inconsistent data sets, volume 4062, 2000.
- [23] P. Guillaume, P. Verboven, S. Vanlanduit, Frequency-domain maximum likelihood identification of modal parameters with confidence intervals, volume 1, 1998, pp. 359–366.
- [24] M. Silva, B. Martinez, E. Figueiredo, J. C. W. A. Costa, Y. Yang, D. Mascareñas, Nonnegative matrix factorization-based blind source separation for full-field and high-resolution modal identification from video, *Journal of Sound and Vibration* 487 (2020) 115586. doi:10.1016/j.jsv.2020.115586.
- [25] X. Fu, K. Huang, N. D. Sidiropoulos, W. K. Ma, Nonnegative matrix factorization for signal and data analytics: Identifiability, algorithms, and applications, *IEEE Signal Processing Magazine* 36 (2019) 59–80. doi:10.1109/MSP.2018.2877582.
- [26] Y. Yang, C. Dorn, Affinity propagation clustering of full-field, high-spatial-dimensional measurements for robust output-only modal identification: A proof-of-concept study, *Journal of Sound and Vibration* 483 (2020) 115473. doi:10.1016/j.jsv.2020.115473.
- [27] B. J. Frey, D. Dueck, Clustering by passing messages between data points, *Science* 315 (2007) 972–976. URL: <https://doi.org/10.1126/science.1136800>. doi:10.1126/science.1136800.
- [28] M. Li, G. Feng, R. Deng, F. Gao, F. Gu, A. D. Ball, Structural vibration mode identification from high-speed camera footages using an adaptive spatial filtering approach, *Mechanical Systems and Signal Processing* 166 (2022) 108422. URL: <https://doi.org/10.1016/j.ymsp.2021.108422>. doi:10.1016/j.ymsp.2021.108422.
- [29] A. A. Ozdemir, S. Gumussoy, Transfer function estimation in system identification toolbox via vector fitting, *IFAC-PapersOnLine* 50 (2017) 6232–6237. doi:10.1016/j.ifacol.2017.08.1026.
- [30] Y. Wang, F. S. Egner, T. Willems, M. Kirchner, W. Desmet, 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 (2022) 108879. URL: <https://doi.org/10.1016/j.ymsp.2022.108879>. doi:10.1016/j.ymsp.2022.108879.
- [31] B. D. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, Morgan Kaufmann Publishers Inc., 1981, pp. 674–679. URL: <https://dl.acm.org/doi/10.5555/1623264.1623280>.
- [32] Ángel J. Molina-Viedma, L. Felipe-Sesé, M. Pastor-Cintas, E. López-Alba, F. A. Díaz, Evaluation of modal identification under base motion excitation using vision techniques, *Mechanical Systems and Signal Processing* 179 (2022) 109405. URL: <https://doi.org/10.1016/j.ymsp.2022.109405>. doi:10.1016/j.ymsp.2022.109405.
- [33] B. Peeters, H. V. der Auweraer, P. Guillaume, J. Leuridan, The polymax frequency-domain method: A new standard for modal parameter estimation?, *Shock and Vibration* 11 (2004) 395–409. doi:10.1155/2004/523692.
- [34] D. L. Brown, R. J. Allemang, R. Zimmerman, M. Mergeay, Parameter estimation techniques for modal analysis, 1979. doi:10.4271/790221.
- [35] I. Tomac, J. Slavič, Morlet-wave-based modal identification in the time domain, *Mechanical Systems and Signal Processing* 192 (2023) 110243. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0888327023001504>. doi:10.1016/J.YMSSP.2023.110243.
- [36] B. K. Horn, B. G. Schunck, Determining optical flow, *Artificial Intelligence* 17 (1981) 185–203. URL: [https://doi.org/10.1016/0004-3702\(81\)90024-2](https://doi.org/10.1016/0004-3702(81)90024-2). doi:10.1016/0004-3702(81)90024-2.

- [37] J. Slavič, M. Boltežar, Damping identification with the morlet-wave, *Mechanical Systems and Signal Processing* 25 (2011) 1632–1645. URL: <https://doi.org/10.1016/j.ymsp.2011.01.008>. doi:10.1016/j.ymsp.2011.01.008.
- [38] S. Mallat, *A Wavelet Tour Of Signal Processing*, 2nd ed., Academic Press, 1999.
- [39] R. Szeliski, *Computer Vision*, Springer International Publishing, 2022. URL: <https://szeliski.org/Book/>. doi:10.1007/978-3-030-34372-9.
- [40] R. C. Gonzalez, R. E. Woods, *Digital Image Processing*, 4 ed., Pearson, 2018.
- [41] J. Manin, S. A. Skeen, L. M. Pickett, Performance comparison of state-of-the-art high-speed video cameras for scientific applications, *Optical Engineering* 57 (2018) 1. URL: <https://doi.org/10.1117/1.OE.57.12.124105>. doi:10.1117/1.OE.57.12.124105.
- [42] R. Costantini, S. Susstrunk, *Virtual sensor design*, 2004, p. 408. doi:10.1117/12.525704.
- [43] W. J. Staszewski, Identification of damping in mdof systems using time-scale decomposition, *Journal of sound and vibration* 203 (1997) 283–305. URL: <https://doi.org/10.1006/jsvi.1996.0864>. doi:10.1006/jsvi.1996.0864.
- [44] J. Slavič, I. Simonovski, M. Boltežar, Damping identification using a continuous wavelet transform, application to real data, *Journal of Sound and Vibration* 262 (2003) 291–307. URL: [https://doi.org/10.1016/S0022-460X\(02\)01032-5](https://doi.org/10.1016/S0022-460X(02)01032-5). doi:10.1016/S0022-460X(02)01032-5.
- [45] I. Tomac, J. Slavič, *ladisk/morletwavemodal: Mwmodal v0.6.3*, 2023. URL: <https://doi.org/10.5281/zenodo.7628043>. doi:10.5281/zenodo.7628043.
- [46] P. L. Reu, E. Toussaint, E. Jones, H. A. Bruck, M. Iadicola, R. Balcaen, D. Z. Turner, T. Siebert, P. Lava, M. Simonsen, Dic challenge: Developing images and guidelines for evaluating accuracy and resolution of 2d analyses, *Experimental Mechanics* 58 (2018) 1067–1099. URL: <https://link.springer.com/article/10.1007/s11340-017-0349-0>. doi:10.1007/s11340-017-0349-0.
- [47] I. Tomac, J. Slavič, Damping identification based on a high-speed camera, *Mechanical Systems and Signal Processing* 166 (2022) 108485. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0888327021008281>. doi:10.1016/j.ymsp.2021.108485.
- [48] K. Zaletelj, D. Gorjup, J. Slavič, *ladisk/pyidi: Release of the version v0.23*, 2020. URL: <https://doi.org/10.5281/zenodo.4017153>. doi:10.5281/zenodo.4017153.
- [49] K. Worden, Data processing and experiment design for the restoring force surface method, part i: integration and differentiation of measured time data, *Mechanical Systems and Signal Processing* 4 (1990) 295–319. URL: [https://doi.org/10.1016/0888-3270\(90\)90010-I](https://doi.org/10.1016/0888-3270(90)90010-I). doi:10.1016/0888-3270(90)90010-I.