Omnidirectional Multi-View High-Speed-Camera-Based Full-Field 3D Modal Identification

Krištof Čufar[a,b], Janko Slavič[a,*]

[a] *University of Ljubljana, Faculty of Mechanical Engineering, Aškerčeva 6, 1000 Ljubljana, Slovenia*
[b] *Domel, d.o.o. Otoki 21, 4228 Železniki, Slovenia*

Cite as:
Krištof Čufar, Janko Slavič; Omnidirectional Multi-View High-Speed-Camera-Based Full-Field 3D Modal Identification, Mechanical Systems and Signal Processing, Voulume 240, November 2025, 113415
https://doi.org/10.1016/j.ymssp.2025.113415

Abstract

Structures with complex geometries often exhibit elaborate spatial responses to a dynamic excitation. Traditional point-wise vibration measurement techniques, such as accelerometer measurements, provide reliable results usually lacking spatial resolution. In contrast, image-based displacement identification methods provide full-field non-contact measurement capabilities at the cost of a lower dynamic range. Furthermore, 3D digital image correlation is able to reconstruct the geometry and spatial displacements of structures, but require multi-camera setups to work. In the case of omnidirectional experimental modal analysis, frequency-domain stitching is required.

This research builds on the recently introduced, multi-view, frequency-domain-triangulation method, that provides a framework for the extraction of full-field 3D operating deflection shapes with a high dynamic range. The existing method is based on sequential high-speed recordings of the vibrating structure made from different views; however, to work it requires prior knowledge of the structure's geometry.

In this research, following an initial camera calibration, a 3D surface mesh is extracted using a photogrammetric geometry reconstruction approach. Displacements of the vibrating object are then recorded and extracted from a large number of views and full-field 3D operating deflection shapes are extracted using frequency-domain triangulation. In the final step, the deflection shapes are magnified and mapped to the reconstructed 3D surface mesh to visualize the vibrational behavior of the test subject. The results of the introduced method for the extraction and visualization of full-field 3D deflection shapes do not require any prior knowledge regarding the geometric or dynamic properties of the studied object. The considerable over-determination in the frequency domain of measurement data obtained from the large number of viewpoints leads to a larger dynamic range and a better reconstruction.

Keywords: experimental modal analysis; high-speed camera; full-field; omnidirectional; modal magnification.

1. Introduction

To visualize a structure's full-field 3D deflection shapes without prior knowledge of its geometry or dynamic properties, three steps must be taken. Firstly, a 3D model of the structure's geometry

*Corresponding author
Email address: janko.slavic@fs.uni-lj.si (Janko Slavič)

needs to be reconstructed. Secondly, its vibrational behavior needs to be studied and thirdly, a visualization of the results must be performed. A short overview of the existing research in the fields of 3D geometry reconstruction, image-based displacement identification and the visualization of magnified deflections is provided in the following section.

The 3D reconstruction of existing objects is a broad area of research and numerous approaches exist. Roughly, they can be classified into contact and non-contact approaches [1, 2], based on the nature of the initial data acquisition. Here, we focus on the non-contact methods, which can be further split into active-vision methods and passive-vision methods. Active methods require some form of structured light to be projected onto the object to extract the geometry information and include methods such as laser scanning [3], structured light scanning [4] and time-of-flight techniques [5], among others. Rinaldi et al. identified the tension in the cables of a cable-stayed footbridge based on the results of a laser-scanning approach [6] The traditional passive-vision-based method of 3D geometry reconstruction involves 2 steps: camera-pose and sparse-scene reconstruction, using Structure-from-motion (SfM) [7] and dense-scene reconstruction, using the Multi-View Stereo (MVS) approach [8, 9]. Traditional SfM methods perform scene optimization by means of the Bundle Adjustment (BA) algorithm [10], either by considering all the images at once in a global approach [11, 12, 13] or incrementally adding images to an initially reconstructed image pair [14, 15, 16]. On the other hand, MVS algorithms can be roughly classified based on their underlying object models into volumetric [17], mesh-based [18], depth-map-based [19] and point-based [9] approaches. Recently, deep-learning-based approaches to both SfM [20, 21, 22] and MVS [23] methods have been widely researched.

Image-based displacement-identification approaches are becoming a viable alternative to traditional methods in the field of structural dynamics [24, 25, 26]. Point tracking [27, 28, 29, 30], (3D) DIC [31, 32, 33, 34, 35, 36, 37], gradient-based optical flow (OF) methods [38, 39, 40, 41, 42] and phase-based OF motion extraction [43, 44, 45, 46, 47] have all been successfully applied. Single-camera 3D DIC approaches have also been extensively studied with the objective of reducing the high initial hardware investment and removing the difficult synchronization of the stereo camera pair of the conventional 3D DIC approach [48]. Three main approaches to single-camera 3D DIC exist; mirror-based [49, 50], prism-based [51, 52] and diffraction-grating-based [53] methods. Yu et al. demonstrated the applicability of the single-camera 3D DIC approach to vibration measurements by extracting the modal parameters of a cantilever plate [54]. Among the listed methods, 3D DIC is the most widely used and has been shown to produce results that are comparable to the laser Doppler vibrometer approach [55, 56]. The extraction of of high-frequency vibration information from low-frame-rate recordings is also possible through the use of under-sampling approaches [57, 58, 59, 55]. Rinaldi et al. performed operational modal analysis based on high-speed camera recordings of a small-scale frame structure and demonstrated the advantages of image-based methods in a damage-detection application [60]. Patil et al. proposed a dynamic stitching approach to address the field-of-view limitation of classic 3D DIC [61]. Gorjup et al. proposed a frequency-domain-triangulation (FDT) method to extract the full-field operating deflection shapes based on sequential recordings from a single high-speed camera [62]. The accuracy of the displacement measurements increases significantly when more than two views are used in the triangulation [63], and consequently, a very large dynamic range can be achieved with the FDT approach. Recently, Baldini et al. employed the FDT approach as a basis for the reconstruction of the vibration-induced acoustic radiation field [64].

As the frequency of the vibration increases, the amplitudes of the observed structure displacements typically decrease [39]. In the field of structural dynamics, the displacements are commonly invisible to the naked eye and covered beneath the noise floor of the camera's sensitivity. Mo-

tion magnification was introduced in 2005 by Liu et al. as a technique for the extraction and visualization of otherwise invisible motion from video data [65]. Since then, the area has seen extensive research and numerous methods have been developed. Motion-magnification methods can be classified into Lagrangian methods [65] that are based on explicit optical flow computation, and Eulerian methods [66], based on a spatio-temporal decomposition of video data. A special subset of the latter group are the phase-based motion-magnification (PBMM) methods [67, 68], which are currently the most popular among all the motion-magnification approaches. PBMM methods have also been applied to the field of structural dynamics [69, 70]. Recently, Feng et al. [71] extended the Eulerian motion-magnification approach to three dimensions using the Neural Radiance Fields (NeRF) deep-learning method [72]. As an alternative to motion magnification-methods, Čufar et al. proposed a modal-analysis-based approach to extract and visualize motion covered under the noise floor of image-based displacement measurements [73]. A recent overview of motion-magnification methods and their application to vibration measurements can be found in [74].

Existing methods for the extraction of full-field spatial vibration information are limited in their field-of-view and are sensitive to image noise. Additionally, they are computationally intensive and often require some prior knowledge regarding the geometry of the test object. A methodology for the extraction and visualization of full-field 3D operating deflection shapes is presented in this manuscript. Firstly, the need for a priori knowledge of the structure's geometry is addressed through the use of a photogrammetric 3D geometry reconstruction approach. Secondly, displacements are identified from a large number of viewpoints, causing considerable over-determination of the triangulation problem and consequently reducing the effect of image noise. Finally, the numerically efficient frequency domain triangulation method is employed to extract omnidirectional vibration information.

The rest of the study is structured as follows. A short overview of the method's theoretical background is given in Sec. 2. The high-speed camera full-field 3D modal magnification method is presented in Sec. 3. The method is validated by means of two experimental cases in Sec. 4 and, finally, the conclusions are drawn in Sec. 5.

2. Theoretical background

This section presents the key theoretical concepts behind the established methods on which the proposed full-field 3D modal magnification method is based. The calibration of the multi-view imaging system provides the basis for the rest of the method and is discussed first in Sec. 2.1. Secondly, the principles of the Structure-from-Motion and Multi View Stereo photogrammetric reconstruction approaches, used to obtain a dense 3D surface representation of the test subject, are introduced in Sec. 2.2. The 2D DIC approach used to extract image-based displacement information is presented in Sec. 2.3. Finally, in Sec. 2.4, the theoretical background is given for the frequency-domain-triangulation method, used to reconstruct the 3D deflection shapes of the test object.

2.1. Calibration of the imaging system

A calibrated imaging system is required to perform the proposed full-field 3D modal magnification method. The projective transformation of 3D object points to their 2D image representations is governed by the projective matrix $\mathbf{P}$, commonly referred to as the camera matrix [75]. It can be expanded into:

$$\mathbf{P} = \mathbf{K}\,[\mathbf{R}|\boldsymbol{t}] \tag{1}$$

The camera matrix in Eq. (1) consists of the intrinsic and extrinsic camera parameters. The latter are presented by the rotation matrix $\mathbf{R}$ and the translation vector $\boldsymbol{t}$ that define the transformation from the world coordinate system to the camera coordinate system (Fig. 1a)). More precisely, the extrinsic parameters present the position and orientation of the camera in the world coordinate system, commonly referred to as the camera pose. The intrinsic parameters, represented by the camera calibration matrix $\mathbf{K}$, characterize the optical properties of the camera, such as its focal length and principal point. The distortion of the lens was not taken into account in this research, as the high-quality lenses with large focal lengths were used for the data acquisition [75]. The camera calibration matrix defines the projection of 3D points from the camera coordinate system onto the 2D image plane (Fig. 1a)) [75].

If the optical system remains constant for the duration of the experiment, the intrinsic camera parameters need to be calibrated only once. Commonly, a set of images of a chessboard pattern is taken and the locations of the pattern's corners are identified on each image. The camera calibration matrix is then reconstructed based on the relation between the known locations of the corners in the world coordinate system and their 2D image projections [76, 77].

The extrinsic parameters need to be reconstructed each time the relative positions of the camera and the observed scene are changed. The extrinsic camera parameters are reconstructed by solving the Perspective-n-Point (PnP) problem [78], based on the known locations of n points in the world coordinate system and their 2D image representations. If the geometry of the scene is unknown, a planar pattern of ArUco markers can be used [79]. In this case, the extrinsic camera parameters are reconstructed in the reference coordinate system, determined by the ArUco board.

Since the definition of the world coordinate system is arbitrary, it is often aligned with the reference coordinate system of the ArUco board to perform the calibration of the camera pose $[\mathbf{R}|\boldsymbol{t}]$ (Fig. 1b)). Alternatively, if the world coordinate system is defined independently, the transformation between it and the ArUco reference frame $[\mathbf{R}|\boldsymbol{t}]_{\mathrm{Ar}}$ must be considered to determine the coordinates of the markers. In this case, solving the PnP problem yields in the camera pose directly in the world coordinate system $[\mathbf{R}|\boldsymbol{t}]_{\mathrm{w}}$ (Fig. 1b)).

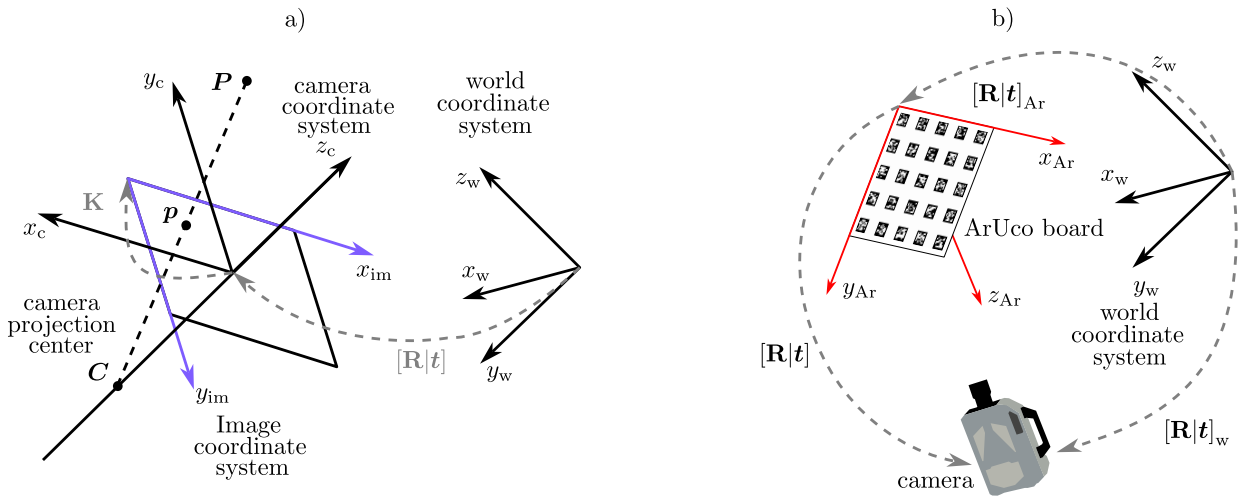

Figure 1: a) Projection of a 3D object point $\boldsymbol{P}$ onto its 2D image representation $\boldsymbol{p}$, b) ArUco board camera calibration

2.2. Photogrammetric Geometry Reconstruction

In this research, the 3D geometry of the test object was reconstructed using a photogrammetric approach. The employed photogrammetric pipeline consists of a Structure-from-Motion stage for a camera-pose estimation and sparse-scene reconstruction, followed by a Multi-View Stereo stage for dense-scene reconstruction. The essential theoretical background to the approach is described in this section.

Structure-from-Motion.
Overlapping images of the object from multiple viewpoints are input along with the camera's intrinsic parameters to initialize the reconstruction. In the first step of the pipeline, feature points need to be identified on every image. Each image is then decomposed into a difference-of-Gaussian pyramid and feature points can then be identified as extrema of the function [80]. Image-intensity gradient orientations in the neighborhood of each feature point are considered to form the Domain-Size Pooling Scale Invariant Feature Transform (DSP-SIFT) feature descriptor [81], employed in the present research. The resulting descriptor is invariant to the feature scale and changes in the feature orientation and illumination.

The feature matching between pairs of images is performed next [82]. For each feature descriptor in the first image of the pair, the closest match is found in the second image. As an exhaustive comparison of all the descriptors would be prohibitively computationally expensive, an approximate nearest-neighbor search is performed based on a KD-tree approach. Repetitive matches are discarded during this step, as a given feature in the first image should also be represented by a single match in the second figure. In the final step of the feature matching, the identified candidates are validated using RANSAC [78].

In the final step of SfM, the camera's extrinsic parameters for each image are reconstructed. Once the feature points have been extracted and matched between the pairs of images, these matches are extended across multiple images to form so-called feature tracks. Each feature track represents a single 3D scene point, as observed from multiple viewpoints. The relation between these estimated 3D points and their 2D image projections is used to solve the Perspective-n-point (PnP) problem [78] and estimate the extrinsic camera parameters for each image. An incremental approach to SfM is employed, where the solution is first obtained for an initial image pair and then incrementally expanded with the other images of the image set [82]. Finally, Bundle Adjustment [10] is performed to optimize both the intrinsic and extrinsic parameters of all the views as well as the locations of the reconstructed 3D scene points. The estimated 3D points used in the PnP problem can be combined into a single point cloud, providing a sparse reconstruction of the observed scene.

Multi-View Stereo.
To achieve a dense reconstruction of the observed scene, depth maps must be estimated. For a reference image, n neighboring images are selected, based on the corresponding camera poses, reconstructed during SfM [82]. The Zero-mean normalized cross-correlation (ZNCC) is computed over a local window, centered on a given pixel and used as a cost function for possible depth values. Then, in a Semi-Global Matching approach [83], the cost is aggregated along several directions across the image to enforce the spatial smoothness. Next, the depth values are identified based on the aggregated cost and a low-resolution depth map is produced. In the final step, the initial depth map is refined to achieve sub-pixel accuracy [82].

In the final step, a 3D surface mesh is obtained. The reconstructed depth maps are merged and a dense point cloud is extracted. Close points are merged to reduce the complexity of the point

cloud [82]. A 3D Delaunay triangulation is then applied to fill the volume of the point cloud, after which the 3D surface mesh can be extracted. Finally, the surface mesh is filtered and smoothed [82].

2.3. Image-based displacement identification

In this study, a 2D DIC approach is used to extract the rigid translations of the small subsets of pixels $\{\Delta x, \Delta y\}^{\mathrm{T}}$ [84]. As the expected displacements are small, the subset deformation and rotation are expected to be negligible [39]. Constraining the motion estimation to translation only is therefore sensible and additionally leads to a significant reduction in the computational time. The difference between a region of interest in the reference image $f(x, y)$ and each consecutive image $g(x+\Delta x, y+\Delta y)$ in the video stream is minimized using an iterative optimization algorithm. The cost function is given as a sum of the squared distances [85]:

$$c_{SSD}(\Delta x, \Delta y) = \sum \left(f(x, y) - g(x + \Delta x, y + \Delta y)\right)^2 \tag{2}$$

In the field of structural dynamics, the displacements to be extracted $(\Delta x, \Delta y)$ can be assumed to be small, and the cost function 2 can be approximated with a first-order Taylor series expansion. The least-squares solution to the problem can be obtained as [85]:

$$\begin{Bmatrix} \Delta x \\ \Delta y \end{Bmatrix} = - \begin{bmatrix} \sum g_x^2 & \sum g_x g_y \\ \sum g_x g_y & \sum g_y^2 \end{bmatrix}^{-1} \begin{Bmatrix} \sum g_x (g - f) \\ \sum g_y (g - f) \end{Bmatrix}, \tag{3}$$

where the first-order derivatives, representing the image-intensity gradients of the displaced image in the x and y directions, are denoted by g_x and g_y. The summations are performed over all the pixels of the subset with a size of $(2M + 1) \times (2M + 1)$ pixels. The reference image $f(x, y)$ is interpolated with bi-variate splines so that sub-pixel displacements can be identified [84]. To ensure successful a displacement identification, the observed surface needs to be covered with a high-contrast pattern, such as a speckle pattern [86].

2.4. Frequency-domain triangulation

In this section the theoretical background to frequency-domain triangulation (FDT) is given. For a more comprehensive explanation of the method, the reader is referred to the original work by Gorjup et al. [62].

The approach is based on the assumption that the observed structure is a linear, time-invariant system. When such a system is subjected to a stationary excitation and is, in response, exhibiting a small harmonic oscillation, its 3D operating deflection shapes can be extracted using a single high-speed camera. Multiple asynchronous recordings of the vibrating structure are obtained by changing the relative positions of the high-speed camera and the structure. As the image sequences at each view point are obtained asynchronously, the relative phase between the views is obtained using a reference accelerometer or force-sensor measurements. The high-speed camera recording and the reference measurement are synchronized using a trigger signal. The trigger signal is generated by the data-acquisition system of the reference measurement and the start of the high-speed camera recording is consequently slightly delayed.

The perspective transformation is used to map 3D object points $\boldsymbol{X} = \{X, Y, Z, 1\}^{\mathrm{T}}$ to their 2D image representations $\boldsymbol{x} = \{x, y, 1\}$ in each view. Here, the points are represented in homogeneous coordinates, as the perspective transform is not linear in Euclidean space. The transformation is given by [75]:

$$z\boldsymbol{x} = \mathbf{P}\boldsymbol{X}, \tag{4}$$

where z is the scaling factor and $\mathbf{P}$ is the camera matrix (Sec. 2.1). Before the FDT method can be performed, the intrinsic parameters of the camera need to be calibrated using a standard chessboard procedure [77]. Care must be taken, that the intrinsic parameters then stay constant for the duration of the experiment. The Perspective-n-Point problem is solved for each view to estimate the rotation matrix and the translation vector from the known locations of n of the 3D object points and their corresponding 2D image representations [78].

After the image-based displacement has been performed for each view, the identified displacement information is transformed into the frequency domain using the Fourier Transform to form frequency-domain images of motion $\hat{\boldsymbol{x}}(\omega_k)$ [62]:

$$\hat{\boldsymbol{x}}(\omega_k) = \frac{1}{z_{\mathrm{REF}}} \mathbf{P} \boldsymbol{X}(\omega_k) \tag{5}$$

The 2D image of the ODS is then identified as the values of the displacement-amplitude spectra in all viewpoints at a given frequency of interest.

Next, multi-view triangulation of the obtained 2D ODS images is performed. To ensure the correct perspective geometry scaling, the identified displacements are first translated to their reference image locations $\boldsymbol{x}_{\mathrm{REF}}$. Consequently, the triangulation in the frequency domain is performed on images of the deflected geometry $\hat{\boldsymbol{x}}_j(\omega_k) = \{\hat{x}_{jk}, \hat{y}_{jk}, 1\}^{\mathrm{T}}$. For each point, two equations can be written [62]:

$$\begin{aligned} \left(\hat{x}_{jk}\boldsymbol{p}_j^3 - \boldsymbol{p}_j^1\right) \boldsymbol{X}\left(\omega_k\right) = 0 \\ \left(\hat{y}_{jk}\boldsymbol{p}_j^3 - \boldsymbol{p}_j^1\right) \boldsymbol{X}\left(\omega_k\right) = 0, \end{aligned} \tag{6}$$

where $\boldsymbol{p}^r$ is the r-th row of the camera matrix, j is the measurement point index and k is the frequency vector index. The homogeneous system of $2n$ equations (n is the number of views, from which a given point is observed) can then be solved by means of a singular value decomposition to obtain the 3D coordinates of $\boldsymbol{X}\left(\omega_k\right)$ [62].

3. Omnidirectional Multi-View High-speed-camera-based full-field 3D modal magnification

The proposed method integrates multiple established techniques from the fields of 3D geometry reconstruction, frequency-domain triangulation and 3D motion reconstruction into a single omnidirectional full-field modal magnification method. Leveraging the numerical efficiency of the frequency-domain-triangulation approach, each displacement measurement point can be observed from a very large number of views (*e.g.*, 20+ views), resulting in a level of measurement redundancy not achievable with other methods. The over-determination of the displacement results in a higher dynamic range and better reconstruction.

A schematic overview of the method's main steps is provided in Fig. 2. In the 1st step of the proposed methodology (Fig. 2a)), the multi-view imaging system is calibrated (Sec. 2.1). An ArUco board with a known location in the world coordinate frame is imaged to reconstruct the pose of the stationary camera. Multiple images of the board in different positions are taken to refine the calibration. In this manuscript, a single high-speed camera is fixed in place and the test object is moved in between the sequential measurements to obtain the desired results. Alternatively, the method could also be performed using a moving camera and stationary object setup or multiple cameras could be used to perform concurrent measurements.

In the 2nd step (Fig. 2b)), the 3D surface model of the test object is reconstructed using the photogrammetric approach, see Sec. 2.2. The test object is imaged from a large number

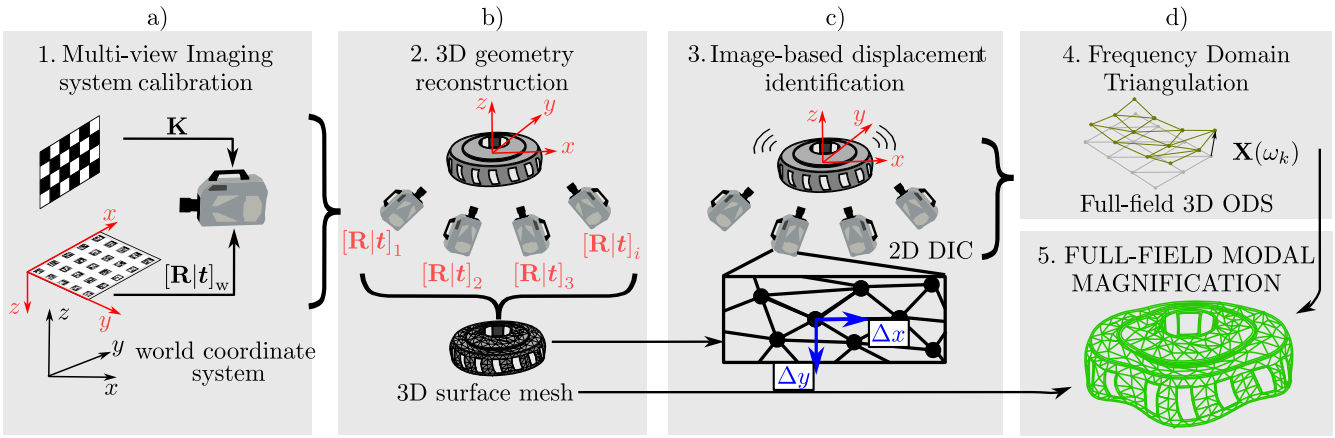

Figure 2: Proposed methodology.

of viewpoints (*e.g.* 144, as shown in Sec. 4.1) while tracking its position in the world coordinate system. Based on the known pose of the stationary camera and the tracked position of the object's coordinate system, the relative camera pose can be reconstructed for each image (analogous to $[\mathbf{R}|\boldsymbol{t}]$ in Fig. 1b)). Supplementary images of the object, acquired from a wider range of viewpoints with an uncalibrated camera, can be integrated into the image set during the Structure-from-Motion stage (Sec. 2.2) to obtain a more accurate and complete reconstruction.

In the 3$^{\text{rd}}$ step (Fig. 2c)), sequential high-speed recordings of the vibrating structure are recorded from a large number of viewpoints (*e.g.*, 24 in the test case, presented in Sec. 4.1) by repositioning the object between recordings. The 3D surface mesh, reconstructed in the 2$^{\text{nd}}$ step (Fig. 2b)), is then reprojected onto the high-speed recordings to determine the displacement measurement locations (Fig. 2c)). Subsequently, the in-plane displacements are identified based on each recording, as described in Sec. 2.3. The still images, required for the 3D geometry reconstruction (step 2), could also be obtained directly from the high-speed recordings (*e.g.* as the average of the recording). However, usually a much larger image set is required to obtain an adequate reconstruction of the object's 3D geometry.

In the 4$^{\text{th}}$ step, the full-field 3D operating deflection shapes are reconstructed using the frequency-domain-triangulation approach, as described in Sec. 2.4 (Fig. 2d)). Similarly to the hybrid-modal-identification approach [87], the natural frequencies of the test object are identified based on the wide-dynamic-range reference accelerometer measurements. Finally, the identified 3D ODS are attributed to the corresponding nodes of the object's 3D surface mesh. The mesh is then deformed in accordance with the scaled 3D ODS to provide a full spherical visualization of the object's vibrations at a specific frequency of oscillation (Fig. 2e)).

The key advantages of the described approach are as follows:

1. No prior knowledge of the object's geometric or dynamic properties is required. This extends the frequency-domain-triangulation approach, where a 3D geometry model must be provided [62].
2. The proposed method enables an omnidirectional full-field analysis fo the object's vibration. No dynamic stitching is required, as in the case of 3D DIC [61].
3. The use of a large number of views, made possible by the numerical efficiency of the frequency-domain-triangulation method, leads to high measurement redundancy and consequently increases the dynamic range of the measurements [63].

4. Experimental demonstration

The proposed full-field modal magnification method was applied to two test objects to showcase its performance. First, an axisymmetric sheet-metal impeller cover of a vacuum-cleaner motor was studied, followed by an experiment on a metal model airplane.

4.1. Impeller cover

An experiment using a sheet-metal impeller cover from a vacuum-cleaner motor (see Fig. 3b)) was conducted to validate the proposed method. The object was expected to exhibit linear behavior and was placed on a soft foam pad to minimize boundary-condition-based non-linearity. The experimental setup is schematically presented in Fig. 3.

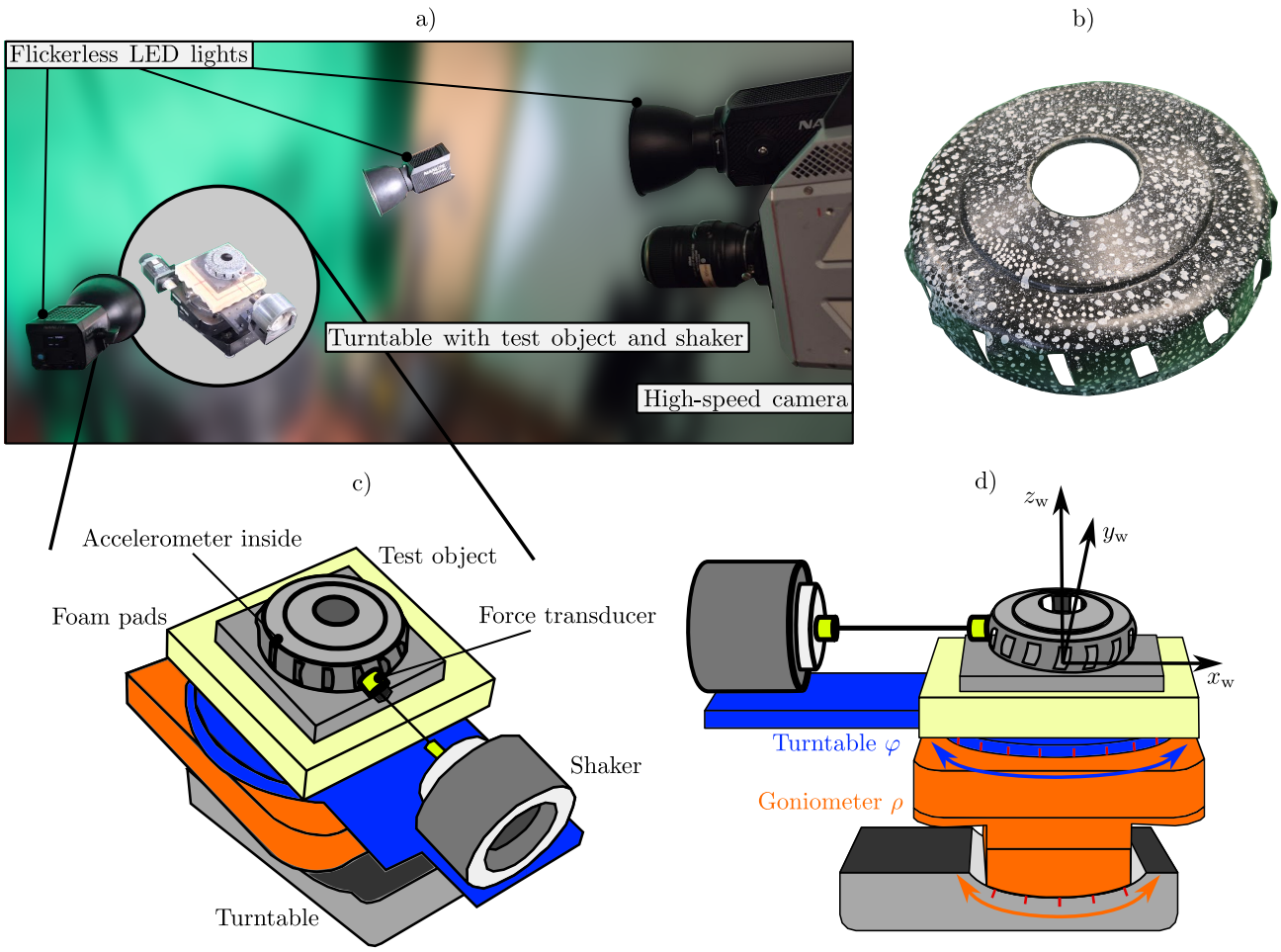

Figure 3: a) Experiment setup for the impeller-cover test case, b) impeller-cover test object, c) schematic close-up representation of the test object, turntable and shaker, d) turntable axis orientations and the position of the world coordinate system.

A single stationary Photron FASTCAM SA-Z high-speed camera was used to capture both the static images, needed for image-based geometry reconstruction, and the high-speed video sequences, on which the displacement identification was performed (Sec. 3). A Sigma 24-70-mm zoom lens with f2.8 was attached to the camera during the experiment. The camera was positioned at a distance of roughly 0.5 m and the focal length of the lens was set at 50 mm. A motorized

and computer-controlled 2-axis turntable was used during all the stages of the experiment for accurate and repeatable repositioning of the impeller cover in the world coordinate system. The configuration of the turntable is shown in Fig. 3c) and d). The first stage, depicted in orange, was used to set the pitch angle ρ of the test object towards the high-speed camera. Then, the second stage of the turntable, depicted in blue, was used to rotate the test object around its own axis by changing the angle φ. The scene was illuminated with three powerful, flickerless LED lights. A single NANLITE Forza 720 served as the primary light source, with two Forza 60 lights providing illumination from the side. (Fig. 3a)).

As discussed in Sec. 3, the camera calibration was performed in the first step of the experiment. Following the calibration of the camera's intrinsic parameters, an ArUco board of 5×5 markers was placed on the turntable to perform the calibration of the multi-view imaging system (Fig. 4a)). A total of 144 images of the board were taken by changing the ρ angle of the turntable from 15° to -7.5° in steps of 7.5° and the φ angle from 0° to 360° in steps of 10°. A mean camera pose in the world coordinate system was calculated in the end, based on the results of all 144 images (Fig. 4b)).

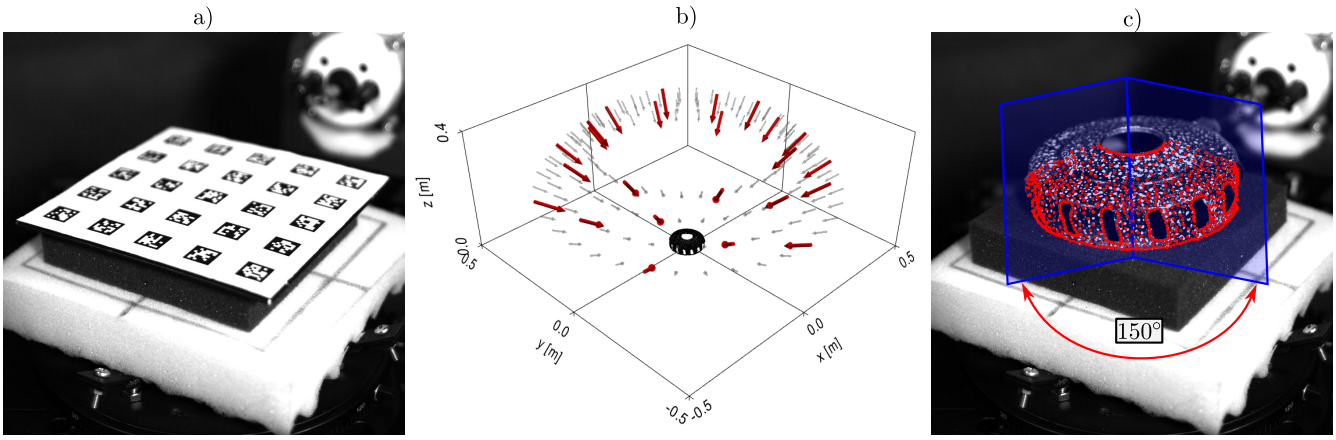

Figure 4: ArUco board used in the multi-view imaging system calibration procedure, b) Camera poses of the static image set (gray and red) and the high-speed recordings (red), c) The displacement measurement locations at $\rho = -7.5°$ and $\varphi = 30°$.

After the calibration of the imaging system, an image-based 3D geometry reconstruction was performed (2^{nd} step). A speckle pattern was applied to the outside surface of the impeller cover to aid both in the process of 3D geometry reconstruction and later during image-based displacement identification. The test object was placed on the turntable and imaged from the same set of viewpoints as the ArUco board in the previous step (one of the obtained images is shown in Fig. 5a)). It should be explained at this point, that the camera pose estimation can be solved directly on the test object images. It is however our experience, that a more accurate and robust solution may be obtained by first performing the pose estimation on ArUco board images. A 3D surface mesh of the impeller cover was then extracted using the Meshroom open-source implementation [82] of the approach, described in Sec. 2.2. The results of the process are shown in Fig. 5b). Some postprocessing was needed to crop the object mesh from the reconstructed background and down-sample the density of the mesh. Additionally, the openings of the impeller cover had to be manually cropped. The result of the postprocessing steps is shown in Fig. 5c), while the wire-frame representation of the final mesh is given in Fig. 5d).

In the 3^{rd} step, vibration measurements were performed on the impeller cover. An electrodynamic shaker was placed on the turntable and attached to the test object via a thin stinger and

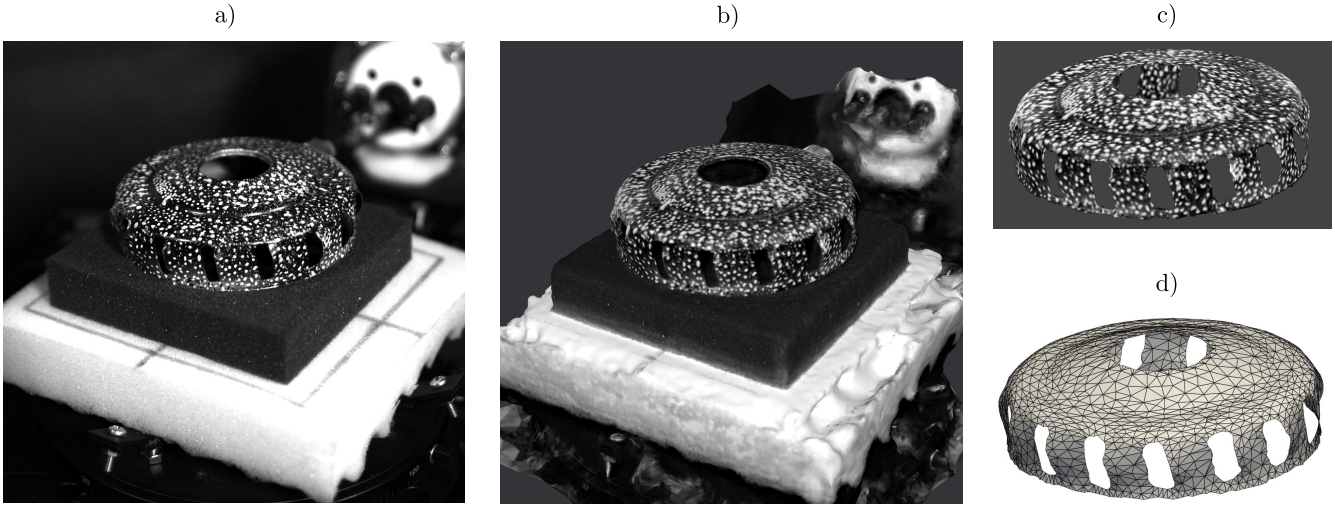

Figure 5: 3D surface-mesh identification process: a) One of the 144 images of the test object, b) the result of photogrammetric geometry reconstruction, c) Cropped surface mesh, d) Wire-frame representation of the experimentally obtained final 3D surface mesh.

a piezoelectric force transducer (Fig. 3). It should be noted that the shaker was attached to the object already during the previous step, before the acquisition of the 3D geometry reconstruction image-set; in this way, the position of the impeller cover on the turntable remained constant for the whole experiment. A pseudo-random signal in the frequency band between 100 and 3000 Hz was used to excite the test object (as implemented in the SDyPy open-source package[1], coauthored by the authors of this manuscript). As the displacement amplitudes of higher-frequency oscillations are typically smaller, the power spectral density (PSD) of the excitation signal was defined as a linear function that increased with the frequency with an amplitude ratio of 1:11 across the frequency band from 100 to 3000 Hz. Consequently, more energy was input into the system at higher frequencies, leading to a more uniform distribution of displacement amplitudes across the observed frequency range.

The camera frame rate was set to 10000 frames per second (FPS) with a 10 μs exposure time. The frame size was set to 896×896 pixels and the bit-depth to 12 bit. A reference accelerometer was attached to the inside of the impeller cover. The sampling frequency of the accelerometer and force transducer measurements was set to 12800 Hz. As the duration of all the measurements was set to 1 s, the frequency resolution in the frequency domain was 1 Hz for the accelerometer, force and high-speed camera measurements. A step trigger signal was used to roughly align the beginning of the camera, force and accelerometer measurements. As the excitation signal was stationary and the system considered was linear and temporally invariant, the slight temporal misalignments were resolved by the transformation of the signals into the frequency domain. There, the detailed phase alignment was performed based on the reference accelerometer measurements, made for each viewpoint (see Sec. 2.4).

To obtain the high-speed recordings, the ρ angle of the turntable (Fig. 3a)) was changed from 15° to -7.5° in steps of 7.5°, while the step of the φ angle was increased by 60°. To achieve a more homogeneous distribution of viewpoints, measurements at $\rho = 0°$ and $\rho = 7.5°$ were offset by 30° along φ. In total, 24 videos of the vibrating structure were recorded (presented

[1]https://github.com/sdypy/sdypy

by red arrows in Fig. 4a)), resulting in 180 GB of video data. The measured force-transducer signal was used to monitor the excitation and ensure that the signal was stationary throughout all 24 measurements. Displacement identification was performed for each view separately. The displacement measurement points were defined by projecting the nodes of the 3D surface mesh onto the high-speed recordings, as described in Sec. 3. Here, the occluded measurement locations had to be filtered out. Given the convex and axisymmetric geometry of the impeller cover, a relatively simple approach was used to filter the occluded points: measurement points were retained only if they fell within a circular sector of the surface mesh, centered around the camera axis. The sector angle was limited to 150° based on a visual inspection of the images of the reprojected points (see Fig. 4c)). The displacement identification was performed using the pyIDI open-source python package implementation of DIC[2] (co-authored by the authors of this manuscript).

Frequency-domain triangulation was then performed to obtain the full-field 3D ODS of the impeller cover. Firstly, all 24 reference high-dynamic-range accelerometer measurements were transformed into the frequency domain where integration into the displacements was performed. An average amplitude spectrum (Fig. 6a)) was then calculated. The peaks in the spectrum were identified to determine the frequencies, at which the spectral ODS were to be reconstructed.

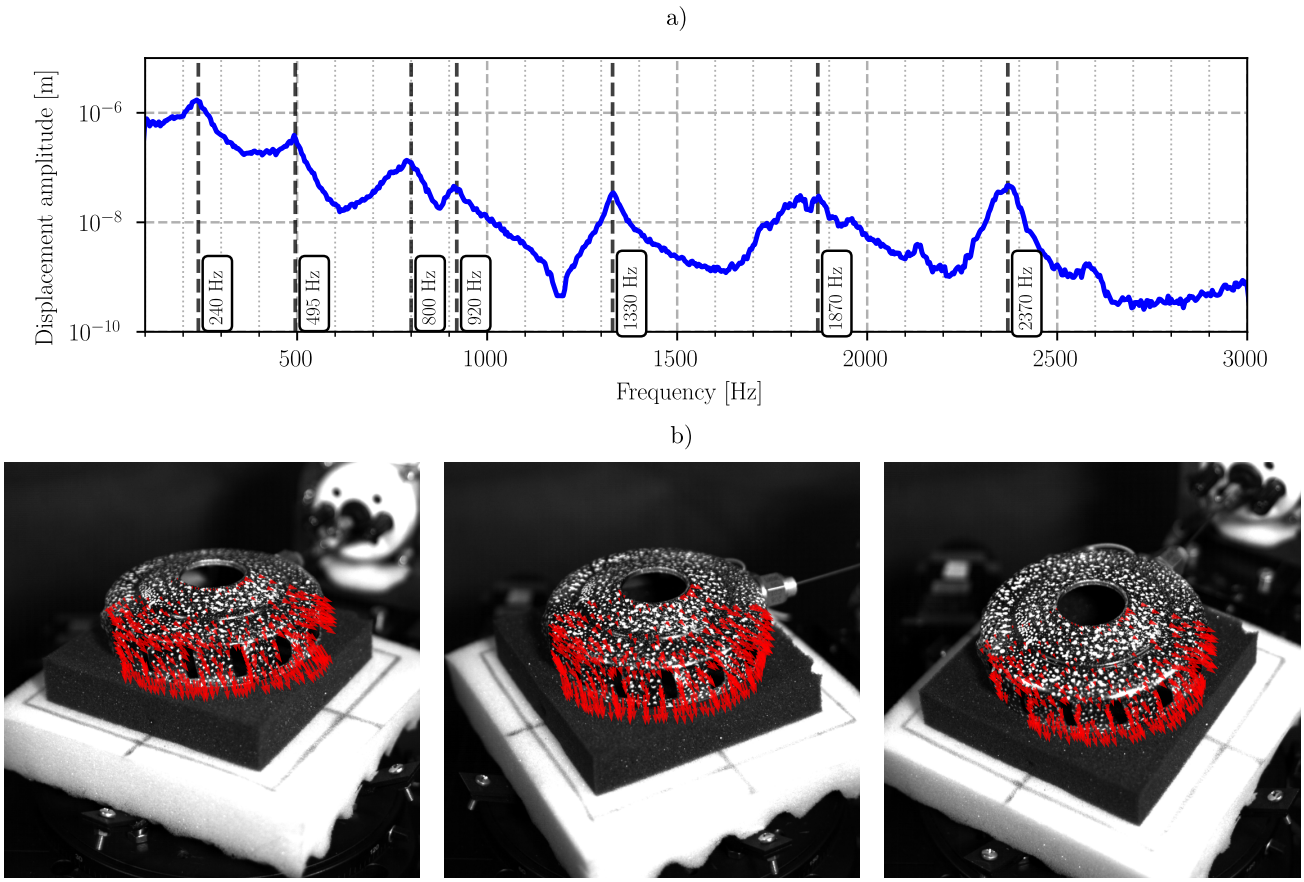

Figure 6: a) Displacement amplitude spectrum, identified based on the reference accelerometer measurements, b) three 2D images of the ODS at 240 Hz.

Secondly, the image-based displacements were transformed into the frequency domain to form

[2]https://github.com/ladisk/pyidi

2D images of the test object's ODS. Fig. 6b) depicts three such images for the first identified ODS at 240 Hz. Displacements were identified from 6 or more viewpoints in 99% of all displacement measurement locations, resulting in a high level of data over-determination. Consequently, the dynamic range of the ODS measurement was increased and a higher quality of reconstruction was achieved. All 2D ODS measurements in a given measurement point were used in the frequency-domain-triangulation process and the full-field 3D ODS. Finally, in the 5^{th} step, the visualization of the results was performed and the 3D surface mesh of the geometry was deformed in accordance with each of the ODS, magnified by an appropriate scaling factor.

The results of the described experiment are shown in Fig. 7. The ODS were identified in 1520 locations (nodes of the reconstructed surface mesh - Fig. 5d)). Seven modes were identified, spread evenly over the entire observed frequency range. As expected, given the radial direction of the excitation, the response of the cylindrical wall of the cover was most prominent. The ODS at 800 and 920 Hz are very similar and were most likely caused by the non-symmetric boundary condition imposed by the shaker stinger. Consequently, only the 800-Hz ODS is shown in Fig. 7c). The maximum displacements at the first mode reached amplitudes of roughly 1 μm, while the displacement amplitude at the sixth mode was two orders of magnitude smaller, despite the ramped excitation signal. The magnification factors used to produce the results, shown in Fig. 7, were calculated so that the displayed maximum deflections have roughly the same amplitude over all six identified modes. While the magnification factors can be chosen arbitrarily, the displayed results show that using the proposed method, displacements with very small amplitudes can be extracted and visualized. The amplitudes of the displacements are given both in metric units and in pixels. The conversion factor of 4.3 pixels/mm was determined as the ratio of the known dimensions of the chessboard calibration pattern in pixels and mm.

Finally, a comment regarding the execution time and computational burden of the described approach is provided. The process of geometry reconstruction varies greatly from case to case. Firstly, the complexity of the geometry dictates, how many views need to be included in the reconstruction. Secondly, smaller details, such as openings in the impeller cover need considerable user interaction in the post-processing stage. For the present test case, roughly 30 minutes was required from the start of the camera calibration process to the generation of the finished surface mesh.

During the displacement identification process, 596 measurement locations were considered on average for each viewpoint and a subset size of 31×31 px was used. The displacement identification was performed on an 6-core AMD Ryzen 5 7600X processor, 32 GiB of DDR6 RAM and an M.2 SSD. On average, each video took 8 minutes to process resulting in a total computation time of 192 minutes for all 24 videos.

The script to perform the post-processing of the displacement data and the FDT required 12.5 s to run and created a memory footprint of 4800 MiB. The DIC results took up most of the memory at 2400 MiB (24 videos at 100 MiB), while the majority of the computation time was required for the transformation of the data into the frequency domain. The FDT itself required less than 1 second to produce all 6 presented ODS and used 1.4 MiB of memory.

The proposed methodology is computationally efficient as displacement identification is performed in the first step of processing the high-speed recordings. Consequently, the amount of data for the subsequent steps is greatly reduced. Similarly, only vibration data at the natural frequencies of the structure is considered, further reducing the computational time of the final triangulation.

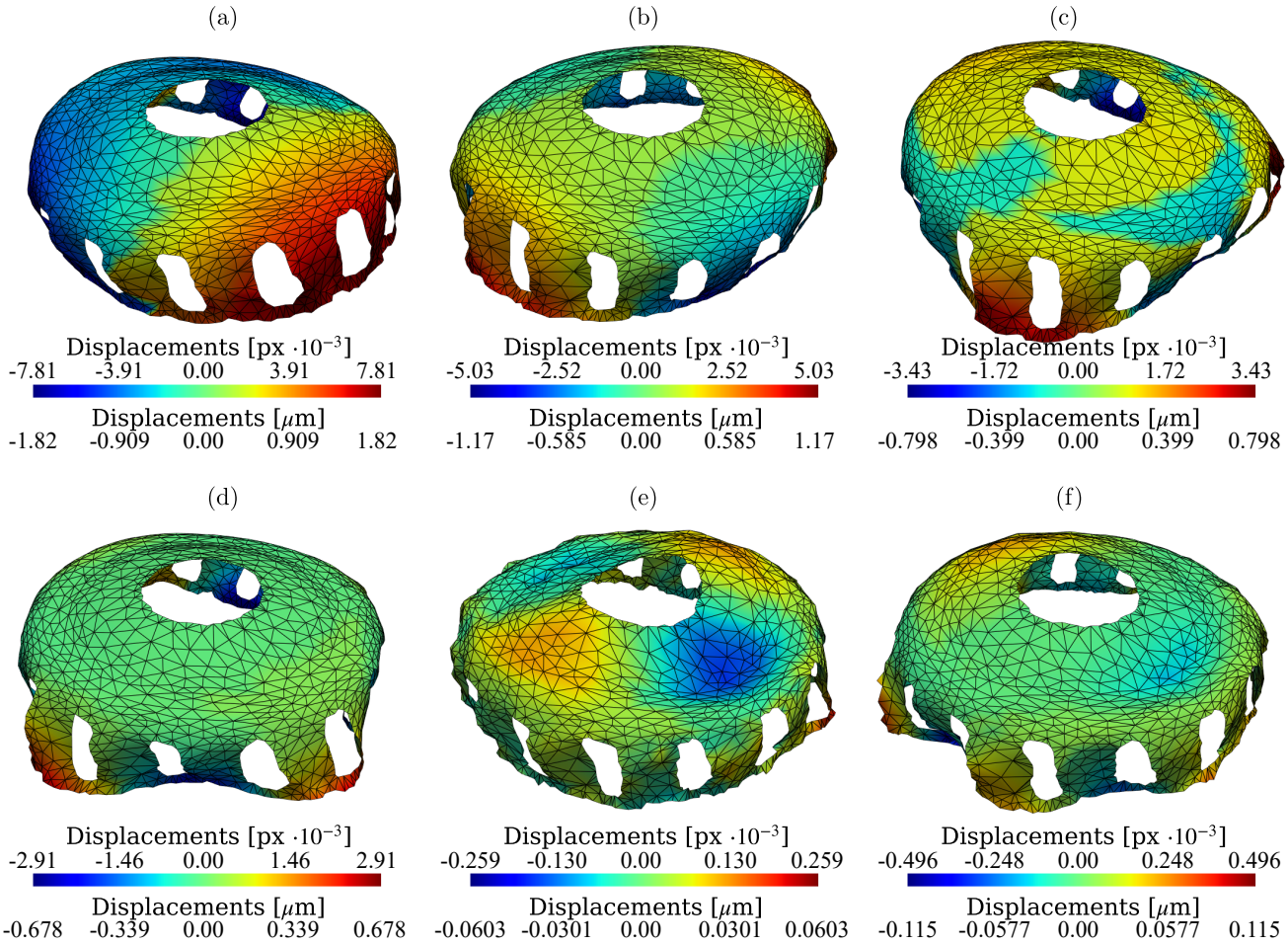

Figure 7: ODS of the impeller cover: a) 240 Hz at 5503 times magnification, b) 495 Hz @ 8547x, c) 800 Hz @ 12528x, d) 1330 Hz @ 14757x, e) 1870 Hz @ 165940x and f) 2370 Hz @ 86632x.

4.2. Airplane model

An additional experiment was performed on a metal airplane model to demonstrate the performance of the proposed method on a more complex geometry and highlight the methods ability to provide an omnidirectional study of an object's dynamic response.

The primary focus of this experiment was the identification and visualization of the flexural deflection shapes of the model's wings. The electrodynamic shaker was repositioned on the turntable to provide excitation in the vertical direction accordingly. The airplane model was attached to the end of the shaker's stinger, lifting it of the turntable and providing clear visibility of its underside.

For image-data acquisition, the camera was positioned to be horizontally aligned with the airplane model to ensure maximum sensitivity for the expected vertical displacements of the wings. During the data acquisition, the ρ angle was set to 15°, 7.5°, 0° and -7.5°, dictated mostly by the limitations of the turntable's movement. The described setup was optimal for displacement identification and additionally enabled the recording of the airplanes underside, resulting in a more complete final result. The experimental setup is show in Fig. 8a).

Again, a speckle pattern was applied to the surface of the test object (Fig. 8b)). The core image-set for 3D geometry reconstruction was obtained using the process, described in Sec. 4.1. To enhance the completeness and detail of the 3D model, an additional set of images was captured using an uncalibrated smartphone camera. These supplementary images improved the coverage of

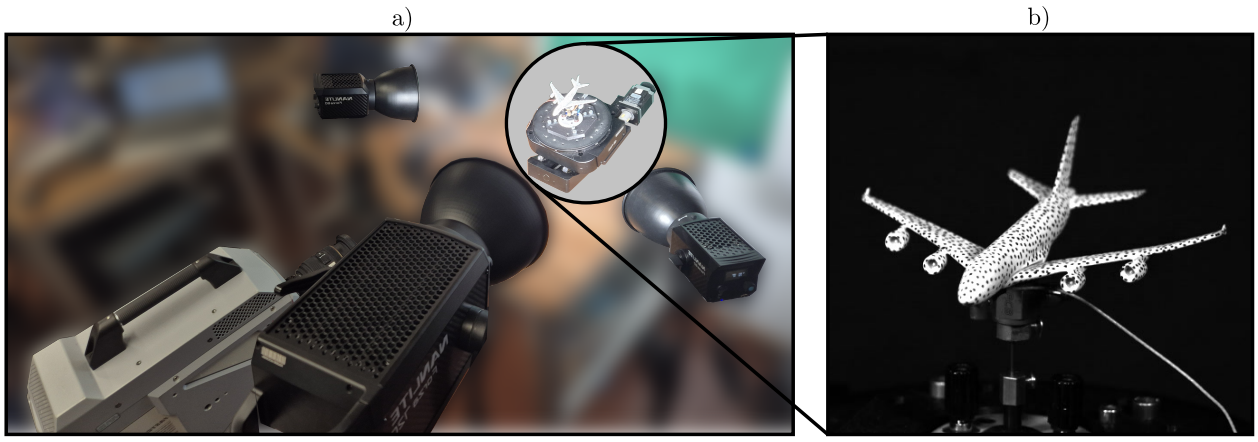

Figure 8: a) Experimental setup for the airplane model test case, b) Airplane model test object.

the areas that were poorly documented using the stationary high-speed camera. The final surface mesh was comprised of 959 nodes.

As in the impeller cover case, a pseudo-random excitation signal was used to excite the object and the data-acquisition parameters remained unchanged. As the airplane's geometry is more complex than the impeller cover in the first case an alternative method was used to determine which displacement measurement locations are visible in a given high-speed recording. An occlusion-culling step was therefore implemented using the OpenCV implementation of ray-tracing [76]. All the mesh nodes were ray-traced to the known camera position of the recording. If a ray intersected with the surface mesh, the corresponding node was considered occluded. The remaining nodes were reprojected onto the recording and resulted in valid displacement measurement points.

The results are depicted in Fig. 9. The first two bending modes of the wings were identified in the observed frequency range. Additionally, a mode of the tail wings was identified (Fig. 9c)). For reference, the undeformed shape of the airplane is also shown faintly on Fig. 9.

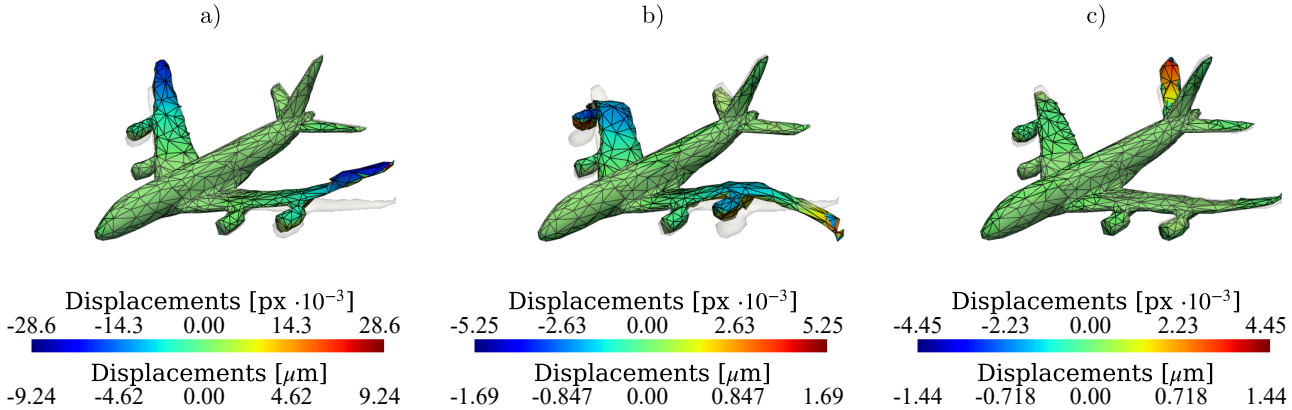

Figure 9: ODS of the airplane model: a) 412 Hz at 2165 times magnification, b) 1393 Hz @ 11805x, c) 1581 Hz @ 13925x

The second flexural mode of the airplanes wings at 1393 Hz (Fig. 9b)) is displayed from three views in Fig. 10 to showcase the completeness of the full-field results, achievable with the proposed methodology.

The distribution of the number of viewpoints from which each ODS measurement location was

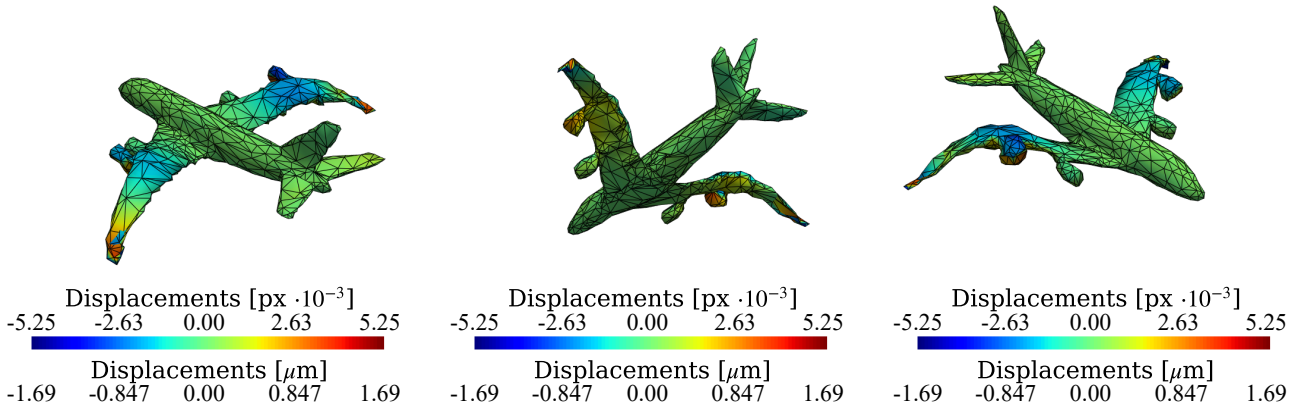

Figure 10: The second flexural mode of the airplane model's wing at 1393 Hz, displayed from three different views.

observed is shown in Fig. 11. In nearly 95% of the 959 locations, the displacements were identified from six or more views. Consequently, a high level of measurement redundancy was achieved, resulting in a very robust and accurate ODS identification.

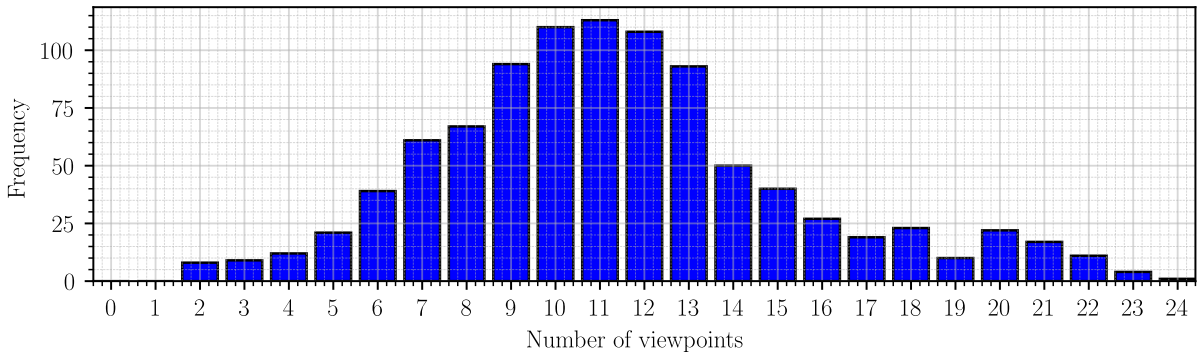

Figure 11: Distribution of measurement locations by number of viewpoints from which they are visible.

The full-field ODS results, presented in Sec. 4.1 and 4.2, are more intuitively visualized in the supplementary video material.

5. Conclusions

The omnidirectional, multi-view, high-speed-camera-based full-field 3D modal magnification method was introduced. The method results in a non-contact extraction and visualization of full-field 3D operating deflection shapes of a structure without any required prior knowledge regarding its geometric or dynamic properties. Following an initial calibration of the imaging system, a set of static images of the test object is obtained and used to produce a 3D surface mesh using a photogrammetric approach. High-speed recordings of the test object's stationary vibrations are then obtained from a large number of viewpoints. 2D image-based displacement identification is then performed for each view using digital image correlation. In the following step, frequency-domain triangulation is performed on the identified 2D displacements to obtain full-field 3D operating deflection shapes of the test object. Finally, the 3D surface mesh is deformed in accordance with the

reconstructed deflection shapes to produce a full-field visualization of the test object's vibrational behavior.

The performance of the proposed methodology was demonstrated on two experimental test cases: a sheet-metal impeller cover of a vacuum-cleaner motor and a metal airplane model. The unknown geometry of the test objects was reconstructed based on a large number of static images, captured with a calibrated camera (*e.g.*, 144 in the case of the impeller cover). In the case of the airplane model, this image-set was extended by additional uncalibrated images to aid with the reconstruction of smaller details. A total of 24 high-speed recordings of the vibrating object were captured for each test case. Most displacement measurement locations were observed from 6 or more view-points (99% in the impeller cover case and 95% in the airplane model case). This resulted in a large over-determination of data for the frequency-domain-triangulation step. Consequently, multiple full-field operating deflection shapes with displacement amplitudes as low as 0.5 of a thousandth of a pixel were extracted and clearly visualized for each test case.

Acknowledgments

The authors acknowledge the partial financial support from the Slovenian Research Agency (research core funding No. P2-0263).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A: The effect of the camera calibration uncertainty on the quality of the 3D operating deflection shapes

The effect of the camera calibration uncertainty on the quality of the 3D operating deflection shapes is quantified in this section. A synthetic Monte Carlo simulation based on a vibrating thin square plate was performed for this purpose.

A flat square plate (100×100 mm) was discretized using a regular grid of 20×20 nodes. The out-of-plane vibrational response of the plate was simulated as a superposition of its first three bending modes (Fig. 12b). The corresponding natural frequencies of the plate were defined at 10, 25 and 60 Hz. There was no phase shift between the modes. The maximum displacement of the plate was scaled to 0.01 mm while the contributions of the three modes were weighted by factors of 1, 0.5 and 0.35 respectively. The synthetic displacement field was generated for all 400 nodes over a 1 second time window at a sampling frequency of 1000 Hz.

A 3×3 array of virtual cameras was then simulated in front of the plate. The cameras were positioned on a sphere with the radius of 0.5 m that was centered at the center of the plate. The spacing of the cameras was 40° both along the azimuth and elevation angles (Fig. 12a)). The cameras were modeled using the pinhole camera mode. The focal length of all cameras was set to 2150 px while the image-size was set to 800×800 px. These parameters were selected to be similar to the setup used in the impeller cover experiment. Once the virtual cameras were defined, the generated displacement field was projected to each of them producing a simulation of DIC displacement extraction. Frequency domain triangulation was performed on this information to produce the ground truth results (Fig. 12b)).

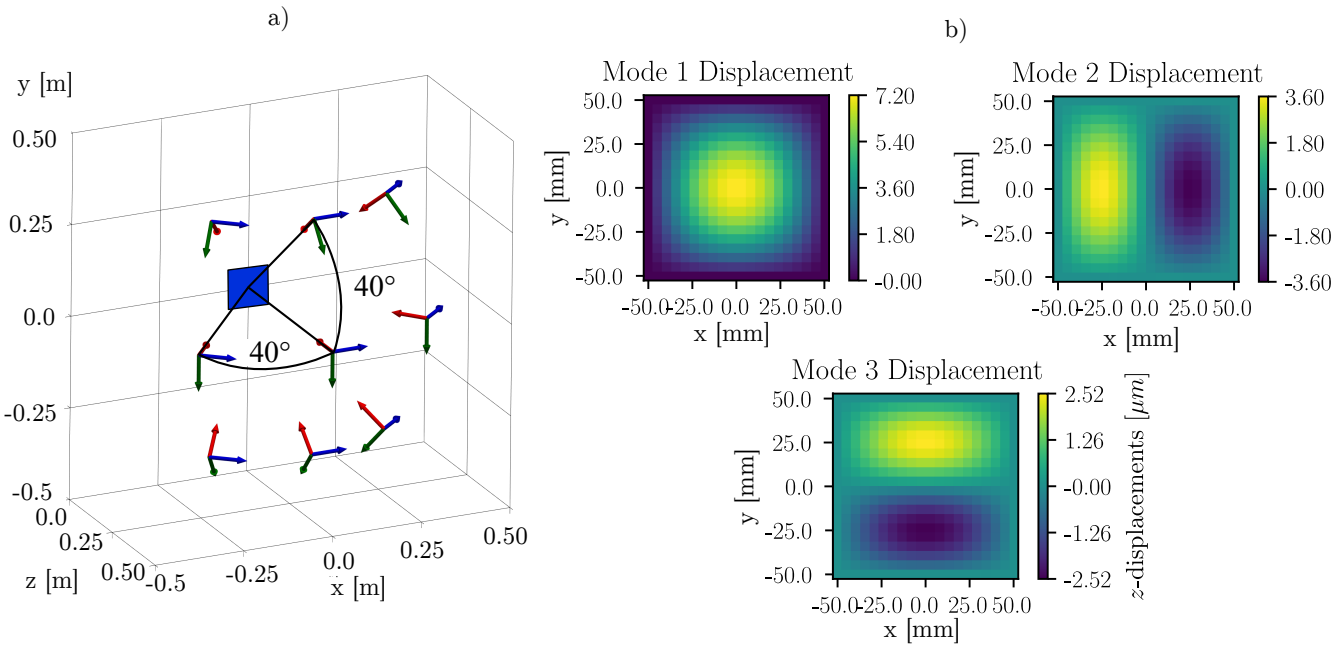

Figure 12: a) Camera setup for the synthetic experiment, b) The ground truth reconstructed mode shapes.

In the next step, the Monte Carlo sensitivity analysis could be performed. The uncertainty of the camera calibration was modeled by introducing Gaussian noise to the intrinsic and extrinsic camera parameters. The standard deviations used to generate the perturbations were as follows: 2px for the focal length, 1px for the principal point coordinates, 0.1 for the quaternion representation of the cameras rotation matrix and 10 mm for the camera position. The perturbed camera parameters were then used to perform the frequency domain triangulation and reconstruct the plates 3D ODS at natural frequencies. 1000 iterations of the Monte Carlo simulation were performed.

The results of the Monte Carlo analysis are shown in Fig. 13. For each of the three ODS, the absolute displacements in the z direction are shown on the top graph. The point indices begin in the bottom left corner of the plate (see Fig. 12b)) and increase along the x axis. For all three modes, the 95% confidence interval is widest at the locations of largest deflections. The histograms of error magnitudes at these locations are displayed on the bottom graphs.

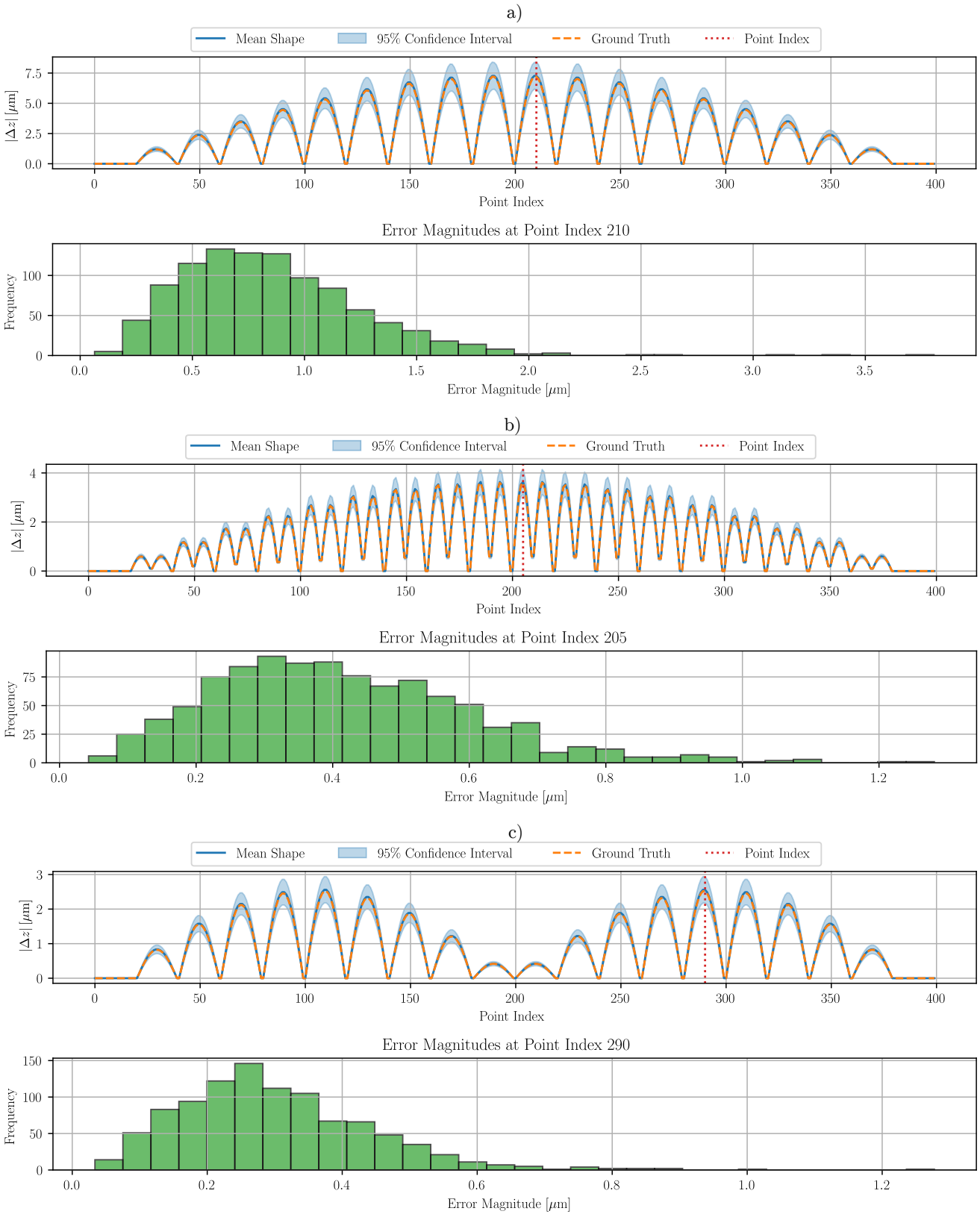

Figure 13: Results of the Monte Carlo analysis: a) First ODS, b) Second ODS, c) Third ODS.

Appendix B: The effect of displacement data over-determination

As cited in the introduction of the present work, Del Sal et al. showed that performing displacement triangulation based on an increasing number of cameras leads to an increase in measurement accuracy [63]. Here, a synthetic experiment was performed to demonstrate the effect of data over-determination on the quality of the final results of the proposed methodology. The same core setup was used, as described in Appendix A (Fig. 12). The azimuth and elevation angle steps were decreased to 20° to generate a denser array of virtual cameras. Four iterations of the synthetic experiment were performed, based on 2, 4, 8 and 24 considered viewpoints respectively. The camera array for each of the iterations is shown on Fig. 14.

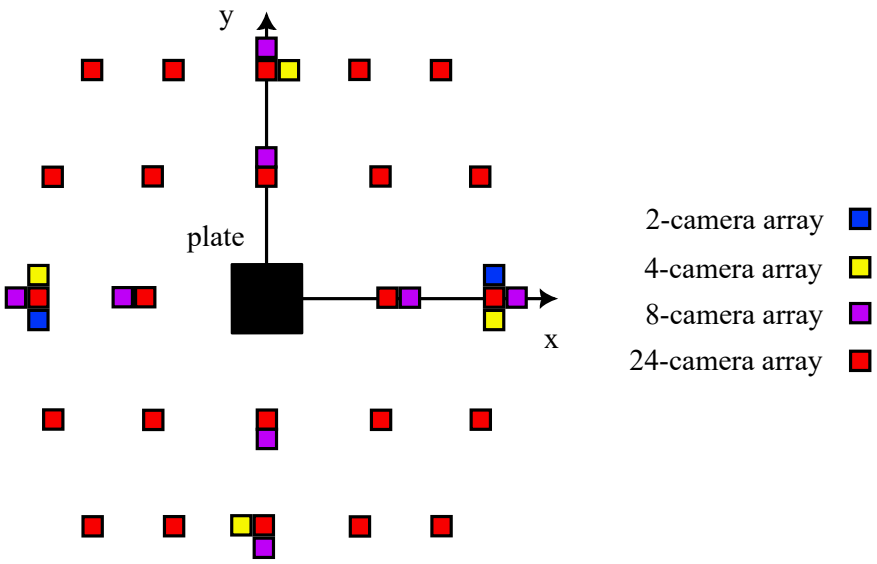

Figure 14: Camera array setup for the synthetic noise study.

A Gaussian noise with the standard deviation of 0.01 px was added to the synthetic image-based displacement measurements of all cameras. Consequently the displacement magnitudes and the noise floor of the measurements were similar to the real-world experiments presented in Sec. 4 of the manuscript.

The Modal Assurance Criterion (MAC) matrices between the ground truth and noise-polluted reconstructed ODS were calculated to visualize the effect of data over-determination on the quality of the final result (Fig. 15).

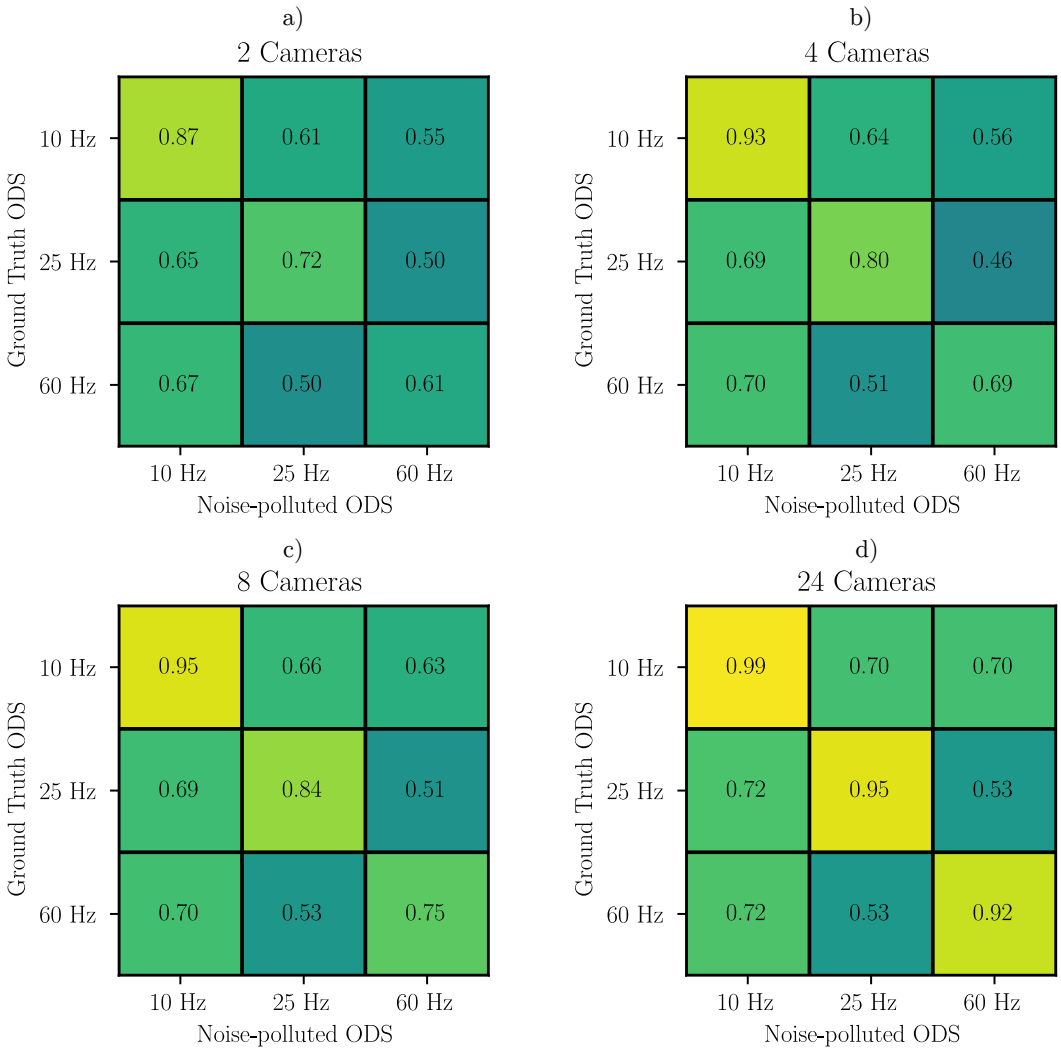

Figure 15: Results of the synthetic noise study, a) 2-camera array, b) 4-camera array, c) 8-camera array and d) 24-camera array.

References

[1] T. Varady, R. R Martin, J. Cox, Reverse engineering of geometric models-an introduction, Computer-Aided Design 29.0 (1997) 255–268. doi:10.1016/S0010-4485(96)00054-1.

[2] Y. Niu, L. Liu, F. Huang, S. Huang, S. Chen, Overview of image-based 3D reconstruction technology, Journal of the European Optical Society-Rapid Publications 20.0 (2024) 18. doi:10.1051/jeos/2024018.

[3] M. A.-B. Ebrahim, 3D Laser Scanners' Techniques Overview, International Journal of Science and Research 4.0 (2015) 323–331. doi:10.21275/SUB158346.

[4] R. Wang, A. C. Law, D. Garcia, S. Yang, Z. Kong, Development of structured light 3D-scanner with high spatial resolution and its applications for additive manufacturing quality assurance, The International Journal of Advanced Manufacturing Technology 117.0 (2021) 845–862. doi:10.1007/s00170-021-07780-2/Published.

[5] Y. Cui, S. Schuon, D. Chan, S. Thrun, C. Theobalt, 3D Shape Scanning with a Time-of-Flight Camera, in: 2010 IEEE Conference on Computer Vision and Pattern Recognition, 2010, pp. 1173–1180. doi:10.1109/CVPR.2010.5540082.

[6] C. Rinaldi, M. Lepidi, F. Potenza, V. Gattulli, Identification of cable tension through physical models and non-contact measurements, Mechanical Systems and Signal Processing 205 (2023). doi:10.1016/j.ymssp.2023.110867.

[7] S. Ullman, The interpretation of structure from motion., in: Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character, 1979, pp. 405–426. doi:10.1098/rspb.1979.0006.

[8] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, R. Szeliski, A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006, pp. 519–528. doi:10.1109/CVPR.2006.19.

[9] Y. Furukawa, J. Ponce, Accurate, dense, and robust multiview stereopsis, IEEE Transactions on Pattern Analysis and Machine Intelligence 32.0 (2010) 1362–1376. doi:10.1109/TPAMI.2009.161.

[10] B. Triggs, P. F. Mclauchlan, R. I. Hartley, A. W. Fitzgibbon, Bundle Adjustment-A Modern Synthesis, in: Vision Algorithms: Theory and Practice: International Workshop on Vision Algorithms, 2000, pp. 298–372. doi:10.1007/3-540-44480-7_21.

[11] D. Crandall, A. Owens, N. Snavely, D. Huttenlocher, Discrete-Continuous Optimization for Large-Scale Structure from Motion, in: CVPR 2011, 2011, pp. 3001–3008. doi:10.1109/CVPR.2011.5995626.

[12] P. Moulon, P. Monasse, R. Marlet, Global fusion of relative motions for robust, accurate and scalable structure from motion, in: Proceedings of the IEEE International Conference on Computer Vision, 2013, pp. 3248–3255. doi:10.1109/ICCV.2013.403.

[13] L. Pan, D. Baráth, M. Pollefeys, J. L. Schönberger, Global Structure-from-Motion Revisited, in: Computer Vision – ECCV 2024, 2024, pp. 58–77. doi:10.48550/arXiv.2407.20219.

[14] N. Snavely, S. M Seitz, R. Szeliski, Photo Tourism: Exploring Photo Collections in 3D, ACM Transactions on Graphics 26.0 (2006) 835–846. doi:10.1145/1141911.1141964.

[15] C. Wu, Towards linear-time incremental structure from motion, in: Proceedings - 2013 International Conference on 3D Vision, 3DV 2013, 2013, pp. 127–134. doi:10.1109/3DV.2013.25.

[16] S. Liu, Y. Gao, T. Zhang, R. Pautrat, J. L. Schönberger, V. Larsson, M. Pollefeys, Robust Incremental Structure-from-Motion with Hybrid Features, in: Computer Vision – ECCV 2024, 2025, pp. 249–269. doi:10.1007/978-3-031-72764-1_15.

[17] S. M. Seitz, C. R. Dyer, Photorealistic Scene Reconstruction by Voxel Coloring, International Journal of Computer Vision 35.0 (1999) 151–173. doi:10.1023/A:1008176507526.

[18] P. Fua, Y. G. Leclerc, Object-Centered Surface Reconstruction: Combining Multi-Image Stereo and Shading, International Journal of Computer Vision 16.0 (1995) 35–56. doi:10.1007/BF01428192.

[19] M. Goesele, B. Curless, S. M. Seitz, Multi-View Stereo Revisited, in: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), 2006, pp. 2402–2409. doi:10.1109/CVPR.2006.199.

[20] C. Tang, P. Tan, BA-Net: Dense Bundle Adjustment Networks, in: International Conference on Learning Representations, 2019, pp. 1–18. doi:10.48550/arXiv.1806.04807.

[21] S. Wang, V. Leroy, Y. Cabon, B. Chidlovskii, J. Revaud, DUSt3R: Geometric 3D Vision Made Easy, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recoginition, 2024, pp. 20697–20709. doi:10.48550/arXiv.2312.14132.

[22] J. Wang, N. Karaev, C. Rupprecht, D. Novotny, Visual Geometry Grounded Deep Structure From Motion, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2024, pp. 21686–21697. doi:10.48550/arXiv.2312.04563.

[23] X. Wang, C. Wang, B. Liu, X. Zhou, L. Zhang, J. Zheng, X. Bai, Multi-view stereo in the Deep Learning Era: A comprehensive revfiew, Displays 70.0 (2021) 102102. doi:10.1016/j.displa.2021.102102.

[24] J. Baqersad, P. Poozesh, C. Niezrecki, P. Avitabile, Photogrammetry and optical methods in structural dynamics – A review, Mechanical Systems and Signal Processing 86.0 (2017) 17–34. doi:10.1016/j.ymssp.2016.02.011.

[25] Y. Cheng, Z. Tian, D. Ning, K. Feng, Z. Li, S. Chauhan, G. Vashishtha, Computer vision-based non-contact structural vibration measurement: Methods, challenges and opportunities, Measurement: Journal of the International Measurement Confederation 243.0 (2025) 116426. doi:10.1016/j.measurement.2024.116426.

[26] Z. Luo, B. Merainani, V. Baltazart, Q. Zhang, M. Döhler, Subpixel motion estimation for video-based target-free vibration monitoring under complex environmental conditions, Mechanical Systems and Signal Processing 226 (2025) 112342. doi:10.1016/j.ymssp.2025.112342.

[27] P. Olaszek, Investigation of the dynamic characteristic of bridge structures using a computer vision method, Measurement 25.0 (1999) 227–236. doi:10.1016/S0263-2241(99)00006-8.

[28] T. G. Ryall, C. S. Fraser, Determination of structural modes of vibration using digital photogrammetry, Journal of Aircraft 39.0 (2002) 114–119. doi:10.2514/2.2903.

[29] A. Khadka, B. Fick, A. Afshar, M. Tavakoli, J. Baqersad, Non-contact vibration monitoring of rotating wind turbines using a semi-autonomous UAV, Mechanical Systems and Signal Processing 138.0 (2020) 106446. doi:10.1016/j.ymssp.2019.106446.

[30] P. Gardonio, G. Guernieri, E. Turco, L. Dal Bo, R. Rinaldo, A. Fusiello, Reconstruction of the sound radiation field from flexural vibration measurements with multiple cameras, Mechanical Systems and Signal Processing 195.0 (2023) 110289. doi:10.1016/j.ymssp.2023.110289.

[31] T. Siebert, R. Wood, K. Splitthof, High speed image correlation for vibration analysis, Journal of Physics: Conference Series 181.0 (2009) 012064. doi:10.1088/1742-6596/181/1/012064.

[32] M. N. Helfrick, C. Niezrecki, P. Avitabile, T. Schmidt, 3D digital image correlation methods for full-field vibration measurement, Mechanical Systems and Signal Processing 25.0 (2011) 917–927. doi:10.1016/j.ymssp.2010.08.013.

[33] W. Wang, J. E. Mottershead, T. Siebert, A. Pipino, Frequency response functions of shape features from full-field vibration measurements using digital image correlation, Mechanical Systems and Signal Processing 28.0 (2012) 333–347. doi:10.1016/j.ymssp.2011.11.023.

[34] R. Zhu, D. Jiang, Z. Huang, L. Xie, D. Zhang, Q. Fei, Full-field modal identification using reliability-guided frequency-domain-based digital image correlation method based on multi-camera system, Measurement: Journal of the International Measurement Confederation 211.0 (2023) 112567. doi:10.1016/j.measurement.2023.112567.

[35] B. A. Furman, B. D. Hill, J. R. Rigby, J. M. Wagner, R. B. Berke, Sensor synchronized DIC: A robust approach to linear and nonlinear modal analysis using low frame rate cameras, Journal of Sound and Vibration 584.0 (2024) 118478. doi:10.1016/j.jsv.2024.118478.

[36] T. Masmeijer, E. Habtour, K. Zaletelj, J. Slavič, Directional DIC method with automatic feature selection, Mechanical Systems and Signal Processing 224.0 (2025) 112080. doi:10.1016/j.ymssp.2024.112080.

[37] W. Ji, Z. Yu, H. Ma, W. Sun, T. Yang, Hyper-reduction modeling and energy transfer analysis of fluid-transporting series-parallel pipes, International Journal of Mechanical Sciences 287 (2025) 109974. doi:10.1016/j.ijmecsci.2025.109974.

[38] B. D. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, in: 7th International Joint Conference on Artificial Intelligence, 1981, pp. 674–679.

[39] J. Javh, J. Slavič, M. Boltežar, The subpixel resolution of optical-flow-based modal analysis, Mechanical Systems and Signal Processing 88.0 (2017) 89–99. doi:10.1016/j.ymssp.2016.11.009.

[40] 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 182.0 (2023) 109553. doi:10.1016/j.ymssp.2022.109553.

[41] Y. Wang, F. S. Egner, T. Willems, F. Naets, M. Kirchner, Using multi-sine excitation and rigid body motion compensation in randomly sampled camera-based experimental modal analysis to improve SNR, Mechanical Systems and Signal Processing 204.0 (2023) 110763. doi:10.1016/j.ymssp.2023.110763.

[42] I. Tomac, J. Slavič, D. Gorjup, Single-pixel optical-flow-based experimental modal analysis, Mechanical Systems and Signal Processing 202.0 (2023) 110686. doi:10.1016/j.ymssp.2023.110686.

[43] A. Sarrafi, P. Poozesh, Z. Mao, A comparison of computer-vision-based structural dynamics characterizations, in: Conference Proceedings of the Society for Experimental Mechanics Series, 2017, pp. 295–301. doi:10.1007/978-3-319-54858-6_29.

[44] B. Merainani, B. Xiong, V. Baltazart, M. Döhler, J. Dumoulin, Q. Zhang, Subspace-based modal identification and uncertainty quantification from video image flows, Journal of Sound and Vibration 569.0 (2024) 117957. doi:10.1016/j.jsv.2023.117957.

[45] X. Yang, P. Yan, X. Liu, W. Lu, X. Zhang, M. Chen, G. Wang, A novel 3D vibration monitoring method using a monocular camera and phase difference Gradient-based algorithm, Measurement: Journal of the International Measurement Confederation 248.0 (2025) 116909. doi:10.1016/j.measurement.2025.116909.

[46] K. Xie, L. Cheng, A novel two-dimensional Hilbert transform method for structural vibration measurement and modal analysis, Mechanical Systems and Signal Processing 225.0 (2025) 112324. doi:10.1016/j.ymssp.2025.112324.

[47] X. Bai, Q. Zhu, X. Wang, Q. Zhang, Y. Du, Modal-weighted super-sensitive phase optical flow method for structural micro-vibration modal identification, Mechanical Systems and Signal Processing 224 (2025) 112095. doi:10.1016/j.ymssp.2024.112095.

[48] B. Pan, L. Ping Yu, Q. B. Zhang, Review of single-camera stereo-digital image correlation techniques for full-field 3D shape and deformation measurement, Science China Technological Sciences 61.0 (2018) 2–20. doi:10.1007/s11431-017-9090-x.

[49] M. Pankow, B. Justusson, A. M. Waas, Three-dimensional digital image correlation technique using single high-speed camera for measuring large out-of-plane displacements at high framing rates, Applied Optics 49.0 (2010) 3418–3427. doi:10.1364/AO.49.003418.

[50] X. Shao, J. Qu, W. Chen, Single-Camera Three-Dimensional Digital Image Correlation with Enhanced Accuracy Based on Four-View Imaging, Materials 16.0 (2023) 2726. doi:10.3390/ma16072726.

[51] K. Genovese, L. Casaletto, J. A. Rayas, V. Flores, A. Martinez, Stereo-Digital Image Correlation (DIC) measurements with a single camera using a biprism, Optics and Lasers in Engineering 51.0 (2013) 278–285. doi:10.1016/j.optlaseng.2012.10.001.

[52] W. Zhang, M. Xu, T. Yuan, W. Li, X. Dai, Z. Zhang, H. Yun, Y. Wang, Dual-biprism-based single-camera high-speed 3D-digital image correlation for deformation measurement on sandwich structures under low velocity impact, Reviews on Advanced Materials Science 62.0 (2023) 20230136. doi:10.1515/rams-2023-0136.

[53] S. Xia, A. Gdoutou, G. Ravichandran, Diffraction Assisted Image Correlation: A Novel Method for Measuring Three-Dimensional Deformation using Two-Dimensional Digital Image Correlation, Experimental Mechanics 53.0 (2013) 755–765. doi:10.1007/s11340-012-9687-0.

[54] L. Yu, B. Pan, Single-camera high-speed stereo-digital image correlation for full-field vibration measurement, Mechanical Systems and Signal Processing 94.0 (2017) 374–383. doi:10.1016/j.ymssp.2017.03.008.

[55] T. J. Beberniss, D. A. Ehrhardt, High-speed 3D digital image correlation vibration measurement: Recent advancements and noted limitations, Mechanical Systems and Signal Processing 86.0 (2017) 35–48. doi:10.1016/j.ymssp.2016.04.014.

[56] P. L. Reu, D. P. Rohe, L. D. Jacobs, Comparison of DIC and LDV for practical vibration and modal measurements, Mechanical Systems and Signal Processing 86.0 (2017) 2–16. doi:10.1016/j.ymssp.2016.02.006.

[57] D. Mastrodicasa, E. Di Lorenzo, S. Manzato, B. Peeters, P. Guillaume, 3D-DIC full field experimental modal analysis of a demo airplane by using low-speed cameras and a reconstruction approach, Mechanical Systems and Signal Processing 227.0 (2025) 112387. doi:10.1016/j.ymssp.2025.112387.

[58] P. Neri, A. Paoli, A. V. Razionale, C. Santus, Low-speed cameras system for 3D-DIC vibration measurements in the kHz range, Mechanical Systems and Signal Processing 162 (2022) 108040. doi:10.1016/j.ymssp.2021.108040.

[59] P. Neri, A. Paoli, A. V. Razionale, S. Barone, Enhanced subpixel sensitivity in 3D-DIC via Spline-Based correlation map interpolation for vibration measurements, Optics and Laser Technology 188 (2025) 112958. doi:10.1016/j.optlastec.2025.112958.

[60] C. Rinaldi, J. Ciambella, V. Gattulli, Image-based operational modal analysis and damage detection validated in an instrumented small-scale steel frame structure, Mechanical Systems and Signal Processing 168 (2022). doi:10.1016/j.ymssp.2021.108640.

[61] K. Patil, V. Srivastava, J. Baqersad, A multi-view optical technique to obtain mode shapes of structures, Measurement: Journal of the International Measurement Confederation 122.0 (2018) 358–367. doi:10.1016/j.measurement.2018.02.059.

[62] D. Gorjup, J. Slavič, M. Boltežar, Frequency domain triangulation for full-field 3D operating-deflection-shape identification, Mechanical Systems and Signal Processing 133.0 (2019) 106287. doi:10.1016/j.ymssp.2019.106287.

[63] R. Del Sal, L. Dal Bo, E. Turco, A. Fusiello, A. Zanarini, R. Rinaldo, P. Gardonio, Structural vibration measurement with multiple synchronous cameras, Mechanical Systems and Signal Processing 157.0 (2021) 107742. doi:10.1016/j.ymssp.2021.107742.

[64] S. Baldini, G. Guernieri, D. Gorjup, P. Gardonio, J. Slavič, R. Rinaldo, 3D sound radiation reconstruction from camera measurements, Mechanical Systems and Signal Processing 227.0 (2025) 112400. doi:10.1016/j.ymssp.2025.112400.

[65] C. Liu, A. Torralba, W. T. Freeman, F. Durand, E. H. Adelson, Motion Magnification, ACM Transactions on Graphics 24.0 (2005) 519–526. doi:10.1145/1073204.1073223.

[66] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, W. T. Freeman, Eulerian video magnification for revealing subtle changes in the world, ACM Transactions on Graphics 31.0 (2012) 1–8. doi:10.1145/2185520.2185561.

[67] N. Wadhwa, M. Rubinstein, F. Durand, W. T. Freeman, Phase-based video motion processing, ACM Transactions on Graphics 32.0 (2013) 1–10. doi:10.1145/2461912.2461966.

[68] N. Wadhwa, M. Rubinstein, F. Durand, W. T. Freeman, Riesz Pyramids for Fast Phase-Based Video Magnification, in: IEEE International Conference on Computational Photography (ICCP), 2014, pp. 1–10. doi:10.1109/ICCPHOT.2014.6831820.

[69] J. G. Chen, N. Wadhwa, Y. J. Cha, F. Durand, W. T. Freeman, O. Buyukozturk, Modal identification of simple structures with high-speed video using motion magnification, Journal of Sound and Vibration 345.0 (2015) 58–71. doi:10.1016/j.jsv.2015.01.024.

[70] G. Liu, M. Z. Li, Z. Mao, Q. S. Yang, Structural motion estimation via Hilbert transform enhanced phase-based video processing, Mechanical Systems and Signal Processing 166.0 (2022) 108418. doi:10.1016/j.ymssp.2021.108418.

[71] B. Y. Feng, H. Alzayer, M. Rubinstein, W. T. Freeman, J.-B. Huang, 3D Motion Magnification: Visualizing Subtle Motions with Time Varying Radiance Fields, in: International Conference on Computer Vision, 2023, pp. 9803–9812. doi:10.1109/ICCV51070.2023.00902.

[72] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, R. Ng, NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, Communications of the ACM 65.0 (2020) 99–106. doi:10.48550/arXiv.2003.08934.

[73] K. Čufar, J. Slavič, M. Boltežar, Mode-shape magnification in high-speed camera measurements, Mechanical Systems and Signal Processing 213.0 (2024) 111336. doi:10.1016/j.ymssp.2024.111336.

[74] K. Luo, X. Kong, J. Li, J. Hu, L. Deng, Motion magnification for video-based vibration measurement of civil structures: A review, Mechanical Systems and Signal Processing 220.0 (2024) 111681. doi:10.1016/j.ymssp.2024.111681.

[75] R. Hartley, A. Zisserman, Multiple View Geometry in computer vision, 2nd Edition, Cambridge University Press, 2003.

[76] G. Bradski, The OpenCV Library, Dr. Dobb's Journal of Software Tools (2000).

[77] Z. Zhang, A Flexible New Technique for Camera Calibration, Pattern Analysis and Machine Intelligence, IEEE Transactions on pattern analysis and machine intelligence 22.0 (2000) 1330–1334. doi:10.1109/34.888718.

[78] M. A. Fischler, R. C. Bolles, Graphics and Image Processing Random Sample Consensus: A Paradigm for Model Fitting with Applicatlons to Image Analysis and Automated Cartography, Communications of the ACM 24.0 (1981) 381–395. doi:10.1145/358669.358692.

[79] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, M. J. Marín-Jiménez, Automatic generation and detection of highly reliable fiducial markers under occlusion, Pattern Recognition 47.0 (2014) 2280–2292. doi:10.1016/j.patcog.2014.01.005.

[80] D. G. Lowe, Object Recognition from Local Scale-Invariant Features, in: Proceedings of the Seventh IEEE International Conference on Computer Vision, 1990, pp. 1150–1157. doi:10.1109/ICCV.1999.790410.

[81] J. Dong, S. Soatto, Domain-Size Pooling in Local Descriptors: DSP-SIFT, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 5097–5106. doi:10.1109/CVPR.2015.7299145.

[82] C. Griwodz, S. Gasparini, L. Calvet, P. Gurdjos, F. Castan, B. Maujean, G. D. Lillo, Y. Lanthony, AliceVision Meshroom, in: ACM Multimedia Systems Conference, 2021, pp. 241–247. doi:10.1145/3458305.3478443.

[83] H. Hirschmüller, Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, pp. 807–814. doi:10.1109/CVPR.2005.56.

[84] B. Pan, K. Li, W. Tong, Fast, Robust and Accurate Digital Image Correlation Calculation Without Redundant Computations, Experimental Mechanics 53.0 (2013) 1277–1289. doi:10.1007/s11340-013-9717-6.

[85] M. A. Sutton, J.-J. Orteu, H. W. Schreier, Image Correlation for Shape, Motion and Deformation Measurements, 1st Edition, Springer New York, NY, 2009.

[86] Y. L. Dong, B. Pan, A Review of Speckle Pattern Fabrication and Assessment for Digital Image Correlation, Experimental Mechanics 57.0 (2017) 1161–1181. doi:10.1007/s11340-017-0283-1.

[87] J. Javh, J. Slavič, M. Boltežar, High frequency modal identification on noisy high-speed camera data, Mechanical Systems and Signal Processing 98.0 (2018) 344–351. doi:10.1016/j.ymssp.2017.05.008.