



HAL
open science

3D vibrations reconstruction with only one camera

Franck Renaud, Stefania Lo Feudo, Jean-Luc Dion, Adrien Goeller

► **To cite this version:**

Franck Renaud, Stefania Lo Feudo, Jean-Luc Dion, Adrien Goeller. 3D vibrations reconstruction with only one camera. *Mechanical Systems and Signal Processing*, 2022, 162, pp.108032. 10.1016/j.ymssp.2021.108032 . hal-03007332

HAL Id: hal-03007332

<https://hal.science/hal-03007332>

Submitted on 16 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

3D vibrations reconstruction with only one camera

Franck Renaud^a, Stefania Lo Feudo^a, Jean-Luc Dion^a, Adrien Goeller^a

^a*SUPMECA, Laboratoire Quartz - EA7393, 3 rue Fernand Hainaut, 93400 Saint-Ouen*

Abstract

In this paper, we propose a new method to measure the vibrations of a 3D mechanical structure. It requires only a video sequence taken by a single camera and a Finite Element model of the structure. First, the intrinsic and extrinsic parameters of the camera are identified. Then, a collection of numerical deflection shapes and normal modes are projected in the image frame of the camera. This is done thanks to a linearization of the perspective projection model. Comparing the motion of targets seen by the camera and the motion of Finite Element model at the same location gives the time evolution of modes amplitude. Finally the motion of the model is reconstructed in 2D in the image frame and in 3D in the world frame. By following this procedure, vibration amplitude can be magnified applying a scale factor on a modal basis covering a large frequency bandwidth. A [video graphical abstract](#) can be found on youtube, see [1].

Keywords: vibrations magnification / 3D vibrations reconstruction / modal expansion / extrinsic calibration / camera / non-linear projection

1. Introduction

In the framework of experimental vibration analysis of civil and mechanical structures, innovation on instrumentation and measurement techniques represents one of the major challenge of the next decades. From the practical point of view, current methods based on local sensors (laser-Doppler vibrometers, laser displacement sensors, piezoelectric accelerometers . . .) allow to measure only a small and limited number of points, a transient full-field measurement is not allowed, and each sensor requires long distance of wire and long preparatory works. To overcome these problems, one can use high-resolution and middle-speed cameras, as well as a network of cameras, in order to catch instantaneously the vibrational behavior of a whole structure. Field vibration measurements based on cameras and video processing represent therefore a new promising way to perform experimental dynamic analysis. Moreover, several methods such as optical flow, data assimilation, digital image correlation, phase-based motion magnification, can be adopted for the image post-processing, cf. [2, 3, 4, 5].

Prior to vibration analysis, camera calibration has to be carried out in order to determine the pose of the camera with respect to the measured object and to correct lens distortions. For this purpose several camera models exist. In the model used by the authors, the Pinhole Camera Model, camera is represented by a box with an infinity small hole on one side. Rays

of light emitted by the 3D object pass through the hole and are caught by the camera sensor upside-down. In the Thin Lens Model, aperture size is no longer negligible, and light rays are refracted passing through the lens frame. More complex models take into account for the lens thickness and the lens geometry, see [6, 7]. Concerning the evaluation of the distortion parameters, in the so-called self-calibration methods, straight edges in the real world are assumed to stay straight on the image. Prescott & McLean [8] proposed to minimize the distance between each point on a given line and a parametric line equation. Devernay & Faugeras [9] suggest to detect 3D edges on the image through a polygonal approximation, and to choose distortion parameters in such a way to minimize the curvature of these lines. Nonlinear distortion correction method proposed by Ricolfe-Viala & Sánchez-Salmerón [10] need fulfillment of a set of restrictions, which are based on line straightness, parallelism and fixed distance of a set of control point placed on a chessboard patten included into the image.

The pose of the camera is identified through extrinsic calibration. It involves estimation of the rotation matrix between world frame and camera frame, as well as cameras relative position if more than one camera is used. Notably, extrinsic parameters can be obtained through homography properties between a planar surface and its image, cf. Zangh [11]. For the 3D reconstruction, several methods exist. For example, matching points or vanishing points can be used, cf. Guillou et al. [12]. Also, the video acquisition system can be completed with an inertial measurement unit able to rotate around two orthogonal axes, cf. Chenguang et al. [13]. A 3D scene can be also recovered by geometric shapes. Wilczkowiak et al. [14] used parallelepipeds, whereas Fan et al. [15] applied a neural networks to approximate the measured shape with a cloud of small spheres. Applications of these techniques can be found in several fields of computer vision, such as building survey, vehicle and objects detection.

Generally, video-based methods for 3D vibration analysis of mechanical structures are based on stereovision, where at least two cameras are used. Digital Image Correlation (DIC) has been recently used for the full-field vibrations measurement. Beberniss & Ehrhardt [16] used 3D-DIC for measurement of beam vibrations. They highlighted the effectiveness of 3D-DIC for modal testing, as well as the possibility to use cameras as virtual strain gauges for deformation measurements. Nevertheless, a few drawbacks exist. The lighting system may alter the signal frequency content, and a major concern is represented by the lack of a temporal aliasing filter within the correlation algorithms [16]. Moreover, DIC allows a full-field vibration measurement of complex structures only for modal displacements greater than noise amplitude, and stereo acquisitions have to be performed, cf. [17, 18].

Concerning utilization of a single camera for 3D analysis, Pan et al. presented in [19] a review of single camera 3D-DIC methods for full-field shape and displacement measurements. Optical devices such as diffraction grating, bi-prism or set of planar mirrors allow to capture an object from different points of view, thus simulating additional virtual cameras. Yu & Pan [20] adopted a settlement composed by a high-speed camera and four mirrors for the full-field 3D vibration measurement of a rectangular plate. Also Durand-Texte et

al. [21] had validated this method on vibration of planar and non planar-objects. Despite of this method reliability, some drawbacks lie on reduction in the spatial resolution, on the difficulty of mirror setting and on the distortion due to mirror's manufacturing errors.

In this paper, we will present an innovative method for the reconstruction of the 3D vibrations of mechanical structures. It requires only a single camera video sequence and a Finite Element model. The method is applied on the structure shown in Figure 2. The structure is mounted on a dynamical shaker which moves with a sinus at 30 Hz during a dozen of periods and then stops. As a result, the structure is subjected to an overall rigid body motion plus modal vibrations causing multi-axial deflections. At high level of amplitude, we expect the mechanical structure to possibly exhibit non-linear behavior due to micro-slip in the bolted lap joints, cf. Festjens et al. [22, 23, 24] or Peyret et al. [25]. Thus normal modes of eigen frequencies higher than 30 Hz are expected to be excited as well. When the shaker stops moving, the amplitude of the normal modes decreases.

The outer surfaces of the structure are bonded with circular targets, which are detected and tracked with the MSER Region algorithm, cf. Matas et al. [26]. For the intrinsic calibration, a classical Pinhole Camera Model and a polynomial model accounting for radial distortions and decentering distortion are used [27]. Then, an extrinsic camera calibration method based on alignment between edges of the image and of a Finite Element model is proposed. The Finite Element model allows to compute a collection of 3D deflection shapes composed of normal modes. These 3D deflection shapes are transformed into 2D deflection shapes thanks to a linearization of the perspective projection. Comparing the motion of targets seen by the camera and the motion of 2D deflection shapes at target location gives the time evolution of deflection shapes amplitude. Finally the motion of the model is reconstructed in 2D in the image frame and in 3D in the world frame. Starting from the time evolution of deflection shapes amplitude, one could estimate damping parameters.

One of the major innovation of this study consists in the 3D reconstruction by projection on the image plane, not only of a static FE model, but also of a dynamic model under small displacements. Moreover, our 3D reconstruction involves the complex shape of the structure driven by discrete points corresponding to the bonded targets. Actually, these points act as virtual sensors, and the projection of normal modes into the image plan is clearly an example of augmented reality. As a result, this method greatly helps in the understanding of the structure dynamics, residual errors being quantifiable.

Actually, the method proposed here can be seen as an alternative to the motion magnification. Two major differences exist between these two methods. First of all, in the motion magnification the choice of the magnification factor is arbitrary and affects the modal parameters identification [2, 28]. On the contrary, in our study magnification is obtained through a scaling factor and it has only a display effect. Second, motion magnification is generally restricted to some frequency bandwidth, and it does not allow separation of closely-spaced modes [29]. On the other hand, our procedure is based on a truncated modal basis which

can cover a large frequency bandwidth. It is worth noticing that this method could also be used to validate a FE model and the assumption of linearity. Indeed, when excited at a given eigenfrequency, the structure will vibrate according to the deflection shape of a single normal mode. If the model is accurate, the numerical deflection shape of this mode is sufficient to explain the observed target motion of the structures, otherwise several deflection shapes will be required.

The paper is organized as follows. In Section 2, intrinsic and extrinsic camera calibrations are presented. Section 3 presents the projection of deflection shapes in the image frame within linearization. Then, in Section 4 3D motion is reconstructed for each target. Modal deflections are presented in Section 5 and method accuracy is discussed as well. Finally some conclusions are drawn in Section 6.

2. Camera calibration

Let us denote \underline{u} the coordinates of a point of the FE model and \underline{v} the coordinates of a point measured with the camera. Figure 1 presents the perspective projection model with the point O as the pinhole. Let us define some coordinate systems and their notations related to the frames shown in Figure 1:

- The world frame \mathcal{W} has G as origin and $\underline{u}_M^{\mathcal{W}} = \{x_M^{\mathcal{W}}, y_M^{\mathcal{W}}, z_M^{\mathcal{W}}\}^t$ are the 3D coordinates of a point M in \mathcal{W}
- The camera frame \mathcal{C} has O as origin and $\underline{u}_M^{\mathcal{C}} = \{x_M^{\mathcal{C}}, y_M^{\mathcal{C}}, z_M^{\mathcal{C}}\}^t$ are the 3D coordinates of a point M in \mathcal{C}
- The image frame \mathcal{I} has O' as origin and $\underline{u}_{M'}^{\mathcal{I}} = \{x_{M'}^{\mathcal{I}}, y_{M'}^{\mathcal{I}}\}^t$ are the 2D coordinates of a point M' in \mathcal{I}
- The pixel frame \mathcal{P} has A as origin and $\underline{u}_{M''}^{\mathcal{P}} = \{x_{M''}^{\mathcal{P}}, y_{M''}^{\mathcal{P}}\}^t$ are the 2D coordinates of a point M'' in \mathcal{P}

Cameras get pictures in the pixel frame of 3D objects described in the world frame. Let us consider a point M described by $\underline{u}_M^{\mathcal{W}}$ in the world frame. To model a camera, first a rigid transformation is applied to find the coordinates $\underline{u}_M^{\mathcal{C}}$ of M in the camera frame. Then, the pinhole camera model is used to account for the perspective projection and to find the coordinates $\underline{u}_{M'}^{\mathcal{I}}$ of M' in the image frame. Finally the polynomial model is used to account for lens distortions and to find the coordinates $\underline{u}_{M''}^{\mathcal{P}}$ of M'' in the pixel frame. The FE model in the world frame \mathcal{W} and the pictures measured by the camera in the pixel frame \mathcal{P} will be compared in the image frame \mathcal{I} . Equation (1) shows the chain of transformations required to make such a comparison.

$$\begin{array}{lcl}
 \text{FE model} & \underline{u}_M^{\mathcal{W}} \rightarrow \underline{u}_M^{\mathcal{C}} \rightarrow \underline{u}_{M'}^{\mathcal{I}} & \\
 & & \underline{u}_{M'}^{\mathcal{I}} \leftarrow \underline{u}_{M''}^{\mathcal{P}} \quad \text{Camera}
 \end{array} \tag{1}$$

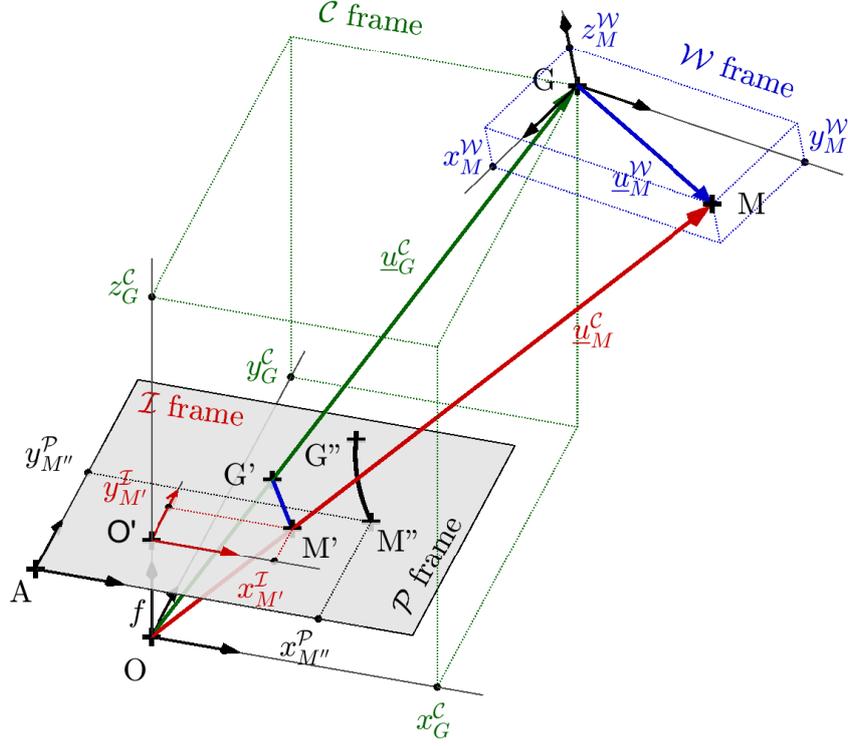


Figure 1: Pose of a camera and perspective projection with the point O as the pinhole.

To transform the coordinates of a point M from world frame to camera frame, a rotation and a translation are applied. The rotation matrix $\underline{R}_{\mathcal{C} \leftarrow \mathcal{W}}$ between the world frame and the camera frame is defined by an unitary quaternion $e_0 + \{e_1, e_2, e_3\}^t$ such that $e_0^2 + e_1^2 + e_2^2 + e_3^2 = 1$, cf. Chou [30]. The expression of the rotation matrix is given by Equation (2). The vector $\underline{u}_G^{\mathcal{C}}$ is the 3D coordinates of G in the camera frame. Then, for a point M with known coordinates $\underline{u}_M^{\mathcal{W}}$ in the world frame, its coordinates $\underline{u}_M^{\mathcal{C}}$ in the camera frame are given by Equation (3). It is worth noticing that the pose of the camera is given by the seven following unknown extrinsic parameters: $x_G^{\mathcal{C}}, y_G^{\mathcal{C}}, z_G^{\mathcal{C}}, e_0, e_1, e_2$ and e_3 .

$$\underline{R}_{\mathcal{C} \leftarrow \mathcal{W}} = \begin{bmatrix} e_0^2 + e_1^2 - e_2^2 - e_3^2 & 2e_1e_2 - 2e_0e_3 & 2e_0e_2 + 2e_1e_3 \\ 2e_0e_3 + 2e_1e_2 & e_0^2 - e_1^2 + e_2^2 - e_3^2 & 2e_2e_3 - 2e_0e_1 \\ 2e_1e_3 - 2e_0e_2 & 2e_0e_1 + 2e_2e_3 & e_0^2 - e_1^2 - e_2^2 + e_3^2 \end{bmatrix} \quad (2)$$

$$\underline{u}_M^{\mathcal{C}} = \underline{R}_{\mathcal{C} \leftarrow \mathcal{W}} \underline{u}_M^{\mathcal{W}} + \underline{u}_G^{\mathcal{C}} \quad (3)$$

Now, let us define f the distance of the image frame to the pinhole. In case of infinity focus, f is also the focal length of the camera. For a point M with known 3D coordinates $\underline{u}_M^{\mathcal{C}} = \{x_M^{\mathcal{C}}, y_M^{\mathcal{C}}, z_M^{\mathcal{C}}\}^t$ in the camera frame, the 2D coordinates $\underline{u}_{M'}^{\mathcal{I}}$ of its image M' in the

image frame are given by Equation (4).

$$\underline{u}_{M'}^{\mathcal{I}} = \frac{f}{z_M^{\mathcal{C}}} \begin{Bmatrix} x_M^{\mathcal{C}} \\ y_M^{\mathcal{C}} \end{Bmatrix} \quad (4)$$

Equations (2), (3) and (4) allow to transform the points from the world frame \mathcal{W} to the image frame \mathcal{I} . Next, let us transform a point M'' from the pixel frame \mathcal{P} into a point M' in the image frame \mathcal{I} . This transformation is applied thanks to the polynomial model in order to account for lens distortions. Parameters K_1 , K_2 and K_3 are needed to model radial distortions and parameters P_1 and P_2 to model the decentering distortions, cf. Halcon manual [27]. Let $\underline{u}_{O'}^{\mathcal{P}} = \{x_{O'}^{\mathcal{P}}, y_{O'}^{\mathcal{P}}\}^t$ denote the 2D coordinates of the optical center in the pixel frame and S_x and S_y denote the pixel size in meters. Equations (5) and (6) allow to transform points from the pixel frame \mathcal{P} to the image frame \mathcal{I} .

$$\begin{cases} x = (x_{M''}^{\mathcal{P}} - x_{O'}^{\mathcal{P}}) \times S_x \\ y = (y_{M''}^{\mathcal{P}} - y_{O'}^{\mathcal{P}}) \times S_y \\ r = \sqrt{x^2 + y^2} \end{cases} \quad (5)$$

$$\underline{u}_{M'}^{\mathcal{I}} = \begin{Bmatrix} x_{M'}^{\mathcal{I}} \\ y_{M'}^{\mathcal{I}} \end{Bmatrix} = \begin{Bmatrix} x(1 + K_1 r^2 + K_2 r^4 + K_3 r^6) + 2P_1 xy + P_2(r^2 + 2x^2) \\ y(1 + K_1 r^2 + K_2 r^4 + K_3 r^6) + P_1(r^2 + 2y^2) + 2P_2 xy \end{Bmatrix} \quad (6)$$

The digital camera used by the authors is a Vieworks VC-12MX-M/C 180 high-speed and high resolution. The hardware and software of the R&D Vision company was used for the measure. The intrinsic parameters of the camera identified with the pattern and software of R&D Vision are: $f = 0.037 \text{ m}$, $K_1 = 65.17 \text{ m.s}^{-2}$, $K_2 = 2.26 \cdot 10^5 \text{ m.s}^{-4}$, $K_3 = -8.36 \cdot 10^8 \text{ m.s}^{-6}$, $P_1 = 0.0102 \text{ m.s}^{-1}$, $P_2 = -0.0016 \text{ m.s}^{-1}$, $S_x = 5.5 \cdot 10^{-6} \text{ m}$, $S_y = 5.5 \cdot 10^{-6} \text{ m}$, $x_{O'}^{\mathcal{P}} = 2173$ pixels and $y_{O'}^{\mathcal{P}} = 248$ pixels. In this study, dynamical behavior of a scale model of an airplane wing has been filmed with only one camera, cf. Figure 2. Several circular targets were bonded on the wing surface for video tracking. The 3D Finite Element (FE) model is depicted in Figure 3. In such case, one possible method for identifying the seven extrinsic parameters is to match points of the world frame to points in the pixel frame in the least square sense. This can be done, for example, thanks to the Matlab Computer Vision Toolbox with the function “extrinsics”.

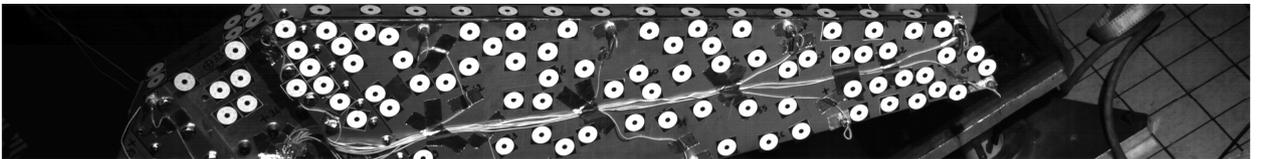


Figure 2: Picture (4096×512) of the mechanical structure filmed with a camera

Here, we present an improvement of this approach for the cases where only few points are available in the picture. Generally speaking, one can easily get by hand corner points on the picture as the intersection of two structure’s edges. Moreover, when corner points lie

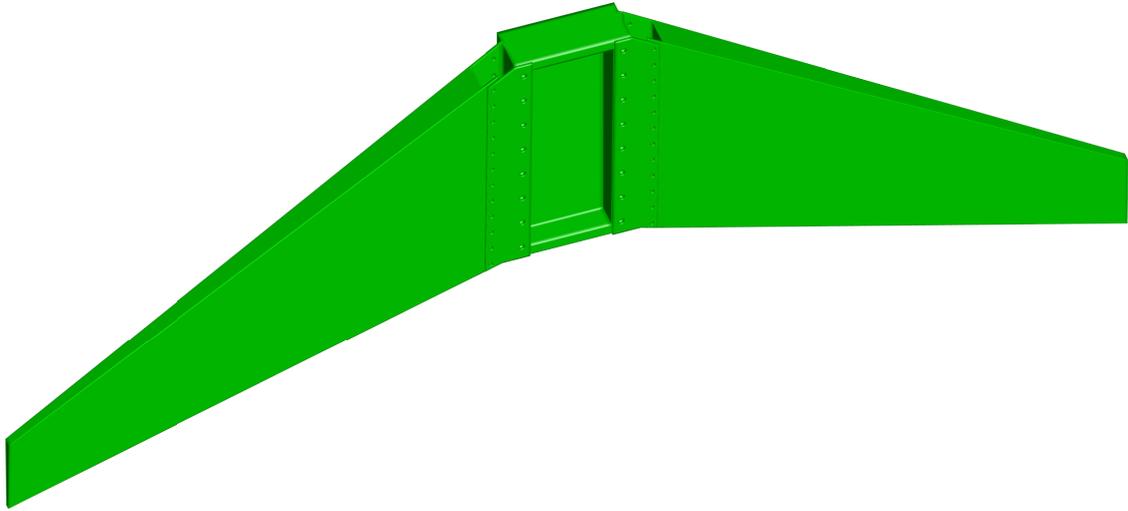


Figure 3: Finite element model of the mechanical structure

outside of the picture one can still rely on the edges. The idea here is to align matching edges coming from the FE model in the world frame \mathcal{W} and the picture in the pixel frame \mathcal{P} . Edges are defined by two points, A and B for the FE model and C and D for the picture, cf. Figure 4. Assuming initial guess for the seven unknown extrinsic parameters: $x_G^C, y_G^C, z_G^C, e_0, e_1, e_2$ and e_3 , all points and lines of interest are transformed into the image frame \mathcal{I} thanks to equations (3), (4), (5) and (6). Let us denote l_A, l_B, l_C and l_D the distance in the image frame between points A, B, C and D and their projection on the other line A', B', C' and D' . When the sum of these length are minimal, the two lines are aligned.

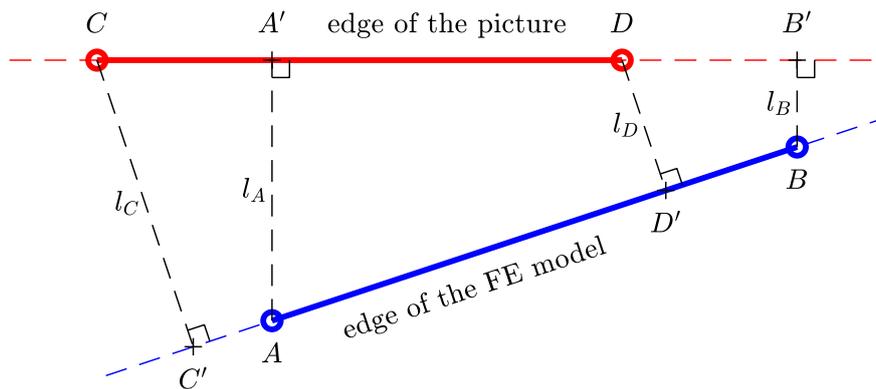


Figure 4: Distance between matching edges on the FE model and on the picture transformed in the image frame \mathcal{I}

The improved objective function w , cf. Equation (7), is a sum of squared lengths: N

lengths l_i between points of interest and the $4 \times M$ lengths l_A , l_B , l_C and l_D between the lines of interests.

$$w = \sqrt{\sum_i^N l_i^2} + \sqrt{\sum_j^M l_A^2 + l_B^2 + l_C^2 + l_D^2} \quad (7)$$

The objective function w is minimized thanks to a classical optimization algorithm. The authors have arbitrary used the Matlab function “fmincon”. The extrinsic parameters minimizing w are: $x_G^{\mathcal{C}} = -0.435$ m, $y_G^{\mathcal{C}} = -0.108$ m, $z_G^{\mathcal{C}} = -3.008$ m, $e_0 = 0.9277$, $e_1 = 0.2596$, $e_2 = 0.0262$ and $e_3 = 0.2670$. Figure 5 shows the picture superimposed with the FE model in the image frame \mathcal{I} after application of the transformation described by Equation (1) with the identified intrinsic and extrinsic parameters.

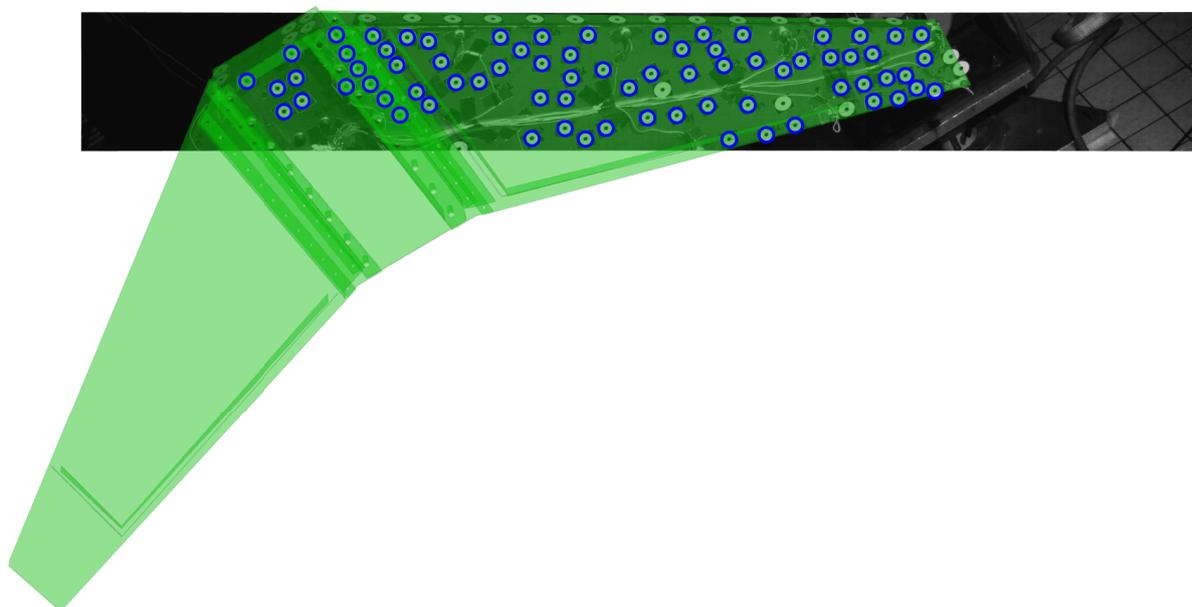


Figure 5: Superimposition of the picture and the FE model. Only $N = 73$ targets were correctly tracked. They are represented with a blue circle.

3. 2D linearized deflection shapes in the image frame

About one hundred targets have been stuck on the mechanical structure, cf. Figure 2. The position of targets are denoted \underline{v}_k when recorded by the camera and they are denoted \underline{u}_k when estimated with the FE model. Target detection and video tracking has been performed through MSER region algorithm, by following the procedure presented in Goeller [31]. Only $N = 73$ targets were correctly tracked by this algorithm. They are shown with a blue circle in Figure 5. As a result, the position $\underline{v}_k^{\mathcal{P}}$ of the center of the target k with $k \in [1..N]$ in the pixel frame \mathcal{P} is known with a subpixel accuracy. Then, the position $\underline{v}_k^{\mathcal{I}}$ of the target k has been computed in the image frame \mathcal{I} by applying Equations (5) and (6).

As explained in the introduction the structure was experiencing vibrations. Eigenvalue analysis has been carried out with the FE model, and the first M normal modes of the structure have been computed. For the work presented in this paper, the chosen deflection shapes $\underline{\psi}_i^{\mathcal{W}}$, with $i \in [1..M]$, are modal ones plus a rigid body motion. One could also choose to use a collection of deflections shapes obtained with different boundary conditions (free, clamped, ...) and even static deflections. Figure 6 shows $\underline{\psi}_1^{\mathcal{W}}$ the rigid body motion and Figures 7, 8, 9 and 10 show $\underline{\psi}_2^{\mathcal{W}}$, $\underline{\psi}_3^{\mathcal{W}}$, $\underline{\psi}_6^{\mathcal{W}}$ and $\underline{\psi}_8^{\mathcal{W}}$. Notably, $\underline{\psi}_i^{\mathcal{W}}$ gives the motion of all the degrees of freedom (dof) of the FE model. Considering $\underline{u}_0^{\mathcal{W}}$ the initial position of the FE mesh and $\underline{\Psi}^{\mathcal{W}}$ the horizontal concatenation of modal deflection shapes $\underline{\psi}_i^{\mathcal{W}}$ with unknown amplitudes $q_i(t)$, we obtain Equation (8). Let $\underline{q}(t) = \{\dots, q_i(t), \dots\}^t$ denote the vector of modal amplitudes.

$$\underline{u}^{\mathcal{W}}(t) = \underline{u}_0^{\mathcal{W}} + \sum_{i=1}^M \underline{\psi}_i^{\mathcal{W}} q_i(t) = \underline{u}_0^{\mathcal{W}} + \underline{\Psi}^{\mathcal{W}} \underline{q}(t) \quad (8)$$

Let us denote $\underline{\phi}_{i|k}^{\mathcal{W}}$ the deflection shape $\underline{\psi}_i^{\mathcal{W}}$ restricted to the target k , wherein $\underline{\phi}_{i|k}^{\mathcal{W}}$ is a 3 rows vector. It has to be pointed out that the nodes of the FE mesh do not perfectly match with the targets position, thus an interpolation has been performed over $\underline{u}_0^{\mathcal{W}}$ and $\underline{\psi}_i^{\mathcal{W}}$. The position of a target in the world frame is the sum of its initial position $\underline{u}_{0|k}^{\mathcal{W}}$ and its modal deflection shapes $\underline{\phi}_{i|k}^{\mathcal{W}}$. Let $\underline{\Phi}_{|k}^{\mathcal{W}} = [\dots, \underline{\phi}_{i|k}^{\mathcal{W}}, \dots]$ denote the 3D modal basis $\underline{\Psi}^{\mathcal{W}}$ restricted to the target k with three rows and M columns. Thus we can assume that the position of the targets in the world frame \mathcal{W} is given by:

$$\underline{u}_k^{\mathcal{W}}(t) = \underline{u}_{0|k}^{\mathcal{W}} + \sum_{i=1}^M \underline{\phi}_{i|k}^{\mathcal{W}} q_i(t) = \underline{u}_{0|k}^{\mathcal{W}} + \underline{\Phi}_{|k}^{\mathcal{W}} \underline{q}(t) \quad (9)$$

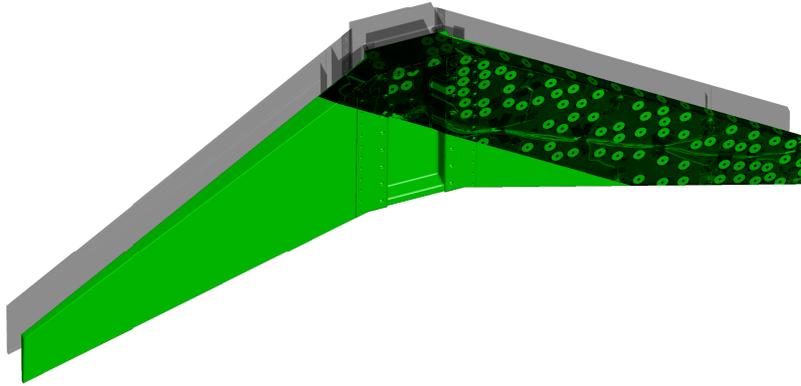


Figure 6: First deflection shape $\underline{\psi}_1^{\mathcal{W}}$. This is the rigid body motion.

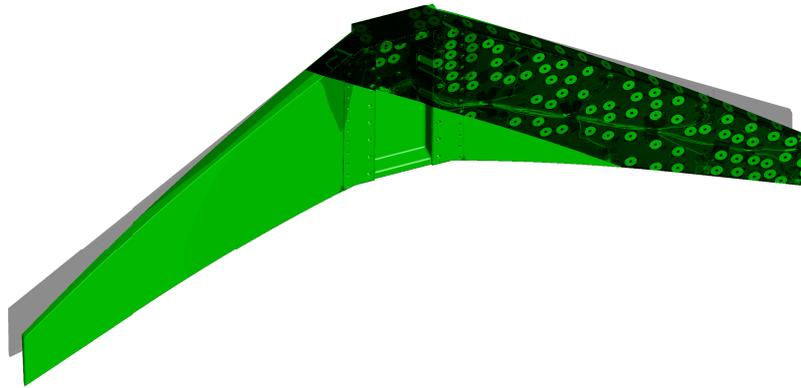


Figure 7: Second deflection shape $\underline{\psi}_2^{\mathcal{W}}$

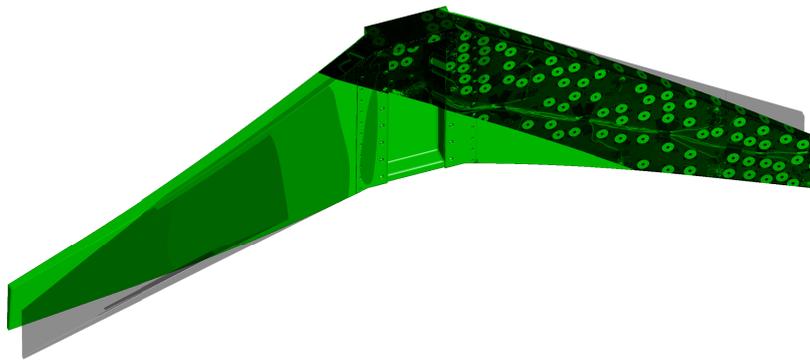


Figure 8: Third deflection shape $\underline{\psi}_3^{\mathcal{W}}$

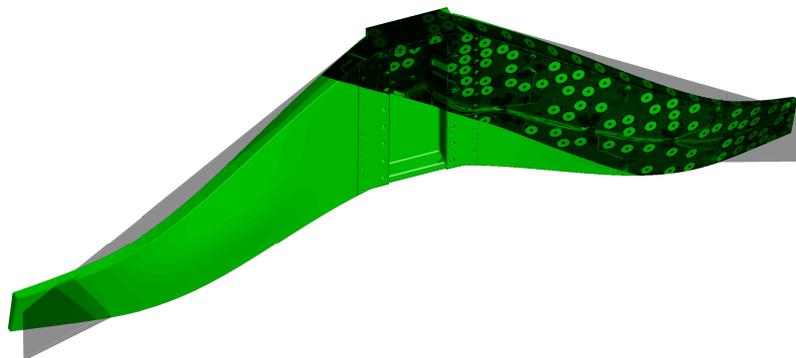


Figure 9: Deflection shape $\underline{\psi}_6^{\mathcal{W}}$

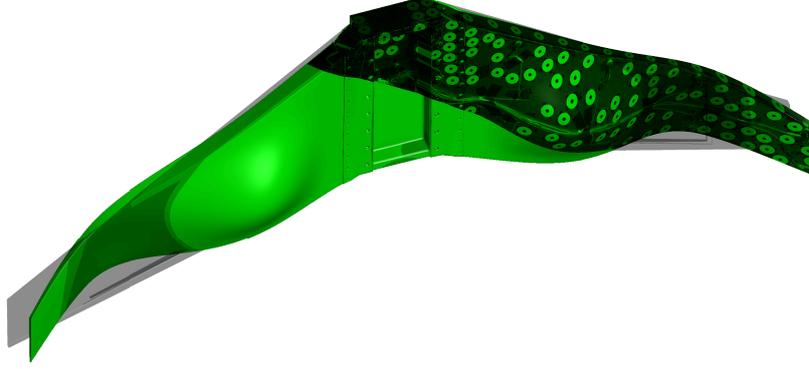


Figure 10: Deflection shape $\underline{\psi}_8^{\mathcal{W}}$

The goal of the work presented here is to identify $\underline{q}(t)$. It is worth noticing that $\underline{q}(t)$ is the same for all the targets. Using Equation (3), the position of the targets in the camera frame \mathcal{C} is given by:

$$\underline{u}_k^{\mathcal{C}}(t) = \underline{u}_{0|k}^{\mathcal{C}} + \underline{\Phi}_{|k}^{\mathcal{C}} \underline{q}(t) \quad (10a)$$

$$\underline{u}_{0|k}^{\mathcal{C}} = \underline{R}_{\mathcal{C} \leftarrow \mathcal{W}} \underline{u}_{0|k}^{\mathcal{W}} + \underline{u}_G^{\mathcal{C}} \quad (10b)$$

$$\underline{\Phi}_{|k}^{\mathcal{C}} = \underline{R}_{\mathcal{C} \leftarrow \mathcal{W}} \underline{\Phi}_{|k}^{\mathcal{W}} \quad (10c)$$

Let's now consider the restriction of the modal shapes of the target k to the x , y and z axis of the camera frame \mathcal{C} . This is represented by $(\underline{\Phi}_{x|k}^{\mathcal{C}})^t$, $(\underline{\Phi}_{y|k}^{\mathcal{C}})^t$ and $(\underline{\Phi}_{z|k}^{\mathcal{C}})^t$ which are respectively the first, second and third rows of the matrix $\underline{\Phi}_{|k}^{\mathcal{C}}$. Using Equation (4), the position of the targets in the image frame \mathcal{I} is given by:

$$\underline{u}_k^{\mathcal{I}}(t) = \frac{f}{z_{0|k}^{\mathcal{C}} + (\underline{\Phi}_{z|k}^{\mathcal{C}})^t \underline{q}(t)} \begin{Bmatrix} x_{0|k}^{\mathcal{C}} + (\underline{\Phi}_{x|k}^{\mathcal{C}})^t \underline{q}(t) \\ y_{0|k}^{\mathcal{C}} + (\underline{\Phi}_{y|k}^{\mathcal{C}})^t \underline{q}(t) \end{Bmatrix} \quad (11)$$

Equation (11) is non linear with respect to $\underline{q}(t)$. However, in the present case, the maximal amplitude of displacement at the end of the wing was about 10^{-2} m, whereas $z_{0|k}^{\mathcal{C}}$ is about 3 m, so that $z_{0|k}^{\mathcal{C}} + (\underline{\Phi}_{z|k}^{\mathcal{C}})^t \underline{q}(t) \simeq z_{0|k}^{\mathcal{C}}$. Thus, the linearisation at first order of Equation (11) is expected to produce an accurate estimation. Its expression is given by:

$$\underline{u}_k^{\mathcal{I}}(t) = \underline{u}_{0|k}^{\mathcal{I}} + \underline{\Phi}_{|k}^{\mathcal{I}} \underline{q}(t) + o(\|\underline{q}(t)\|) \quad (12a)$$

$$\underline{u}_{0|k}^{\mathcal{I}} = \frac{f}{z_{0|k}^{\mathcal{C}}} \begin{Bmatrix} x_{0|k}^{\mathcal{C}} \\ y_{0|k}^{\mathcal{C}} \end{Bmatrix} \quad (12b)$$

$$\underline{\Phi}_{|k}^{\mathcal{I}} = \frac{f}{(z_{0|k}^{\mathcal{C}})^2} \begin{Bmatrix} (\underline{\Phi}_{x|k}^{\mathcal{C}})^t z_{0|k}^{\mathcal{C}} - x_{0|k}^{\mathcal{C}} (\underline{\Phi}_{z|k}^{\mathcal{C}})^t \\ (\underline{\Phi}_{y|k}^{\mathcal{C}})^t z_{0|k}^{\mathcal{C}} - y_{0|k}^{\mathcal{C}} (\underline{\Phi}_{z|k}^{\mathcal{C}})^t \end{Bmatrix} \quad (12c)$$

After concatenation for all targets we get:

$$\underline{u}^{\mathcal{I}}(t) = \underline{u}_0^{\mathcal{I}} + \sum_{i=1}^M \underline{\Phi}_i^{\mathcal{I}} q_i(t) = \underline{u}_0^{\mathcal{I}} + \underline{\underline{\Phi}}^{\mathcal{I}} \underline{q}(t) \quad (13a)$$

$$\underline{u}^{\mathcal{I}}(t) = [\dots, (\underline{u}_k^{\mathcal{I}}(t))^t, \dots]^t \quad (13b)$$

$$\underline{u}_0^{\mathcal{I}} = [\dots, (\underline{u}_{0|k}^{\mathcal{I}})^t, \dots]^t \quad (13c)$$

$$\underline{\underline{\Phi}}^{\mathcal{I}} = [\dots, (\underline{\underline{\Phi}}_{|k}^{\mathcal{I}})^t, \dots]^t \quad (13d)$$

Therefore, for N targets and M normal shapes, the matrix $\underline{\underline{\Phi}}^{\mathcal{I}}$ has $2N$ rows and M columns. $\underline{\underline{\Phi}}^{\mathcal{I}}$ is the horizontal concatenation of targets deflection shapes in the image frame $\underline{\phi}_i^{\mathcal{I}}$. Figure 11 shows the target deflection shape $\underline{\phi}_1^{\mathcal{I}}$ associated with the rigid body motion shown in Figure 6. Figures 12, 13, 14, 15 respectively show the target deflection shapes $\underline{\phi}_2^{\mathcal{I}}$, $\underline{\phi}_3^{\mathcal{I}}$, $\underline{\phi}_6^{\mathcal{I}}$ and $\underline{\phi}_8^{\mathcal{I}}$ associated with the 3D deflection shapes shown in Figures 7, 8, 9 and 10.

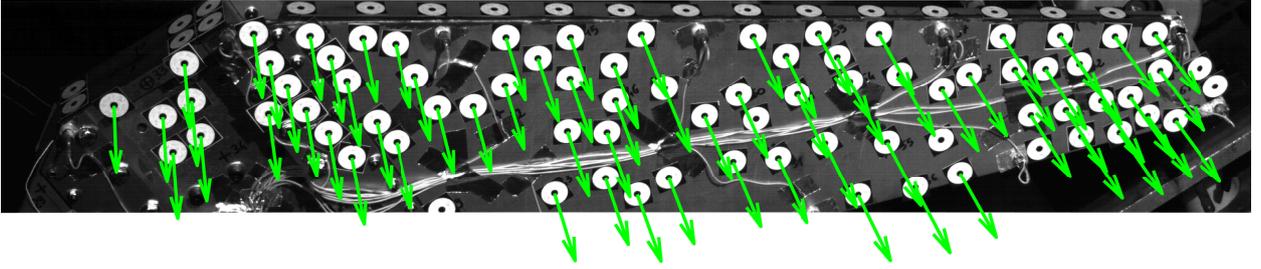


Figure 11: First deflection shape $\underline{\phi}_1^{\mathcal{I}}$. This is the rigid body motion.

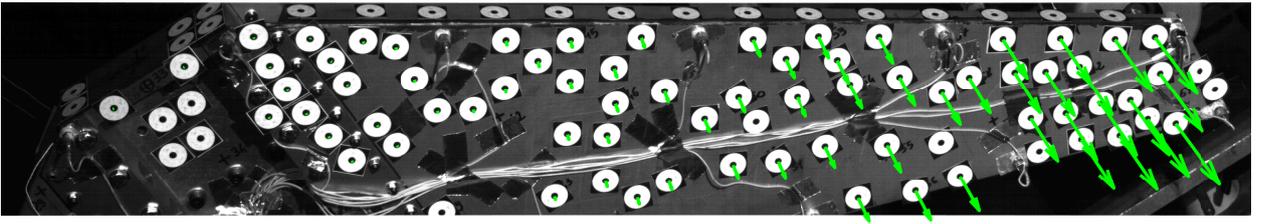


Figure 12: Second deflection shape $\underline{\phi}_2^{\mathcal{I}}$

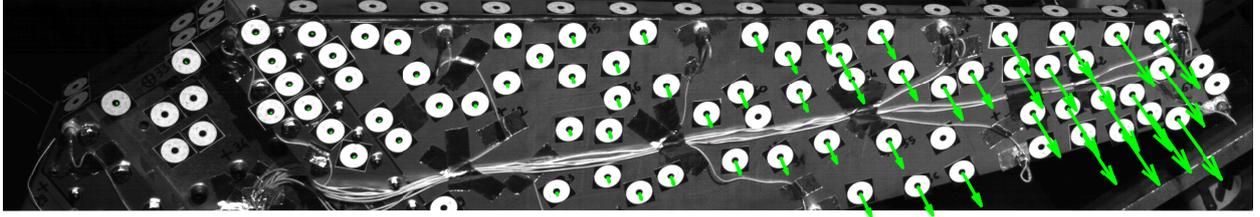


Figure 13: Third deflection shape $\underline{\phi}_3^I$

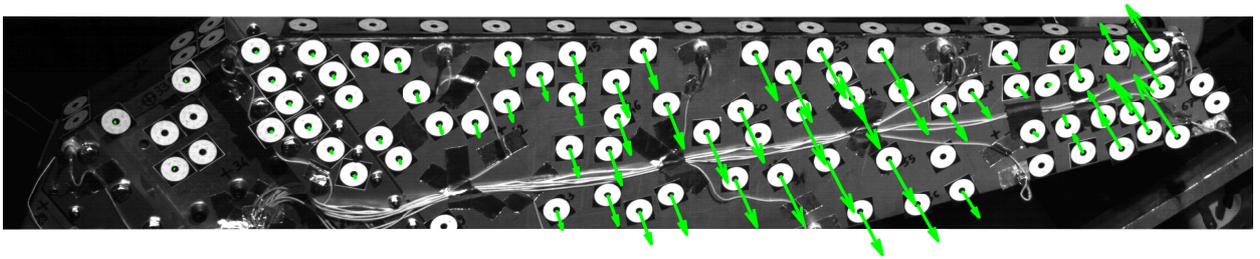


Figure 14: Deflection shape $\underline{\phi}_6^I$

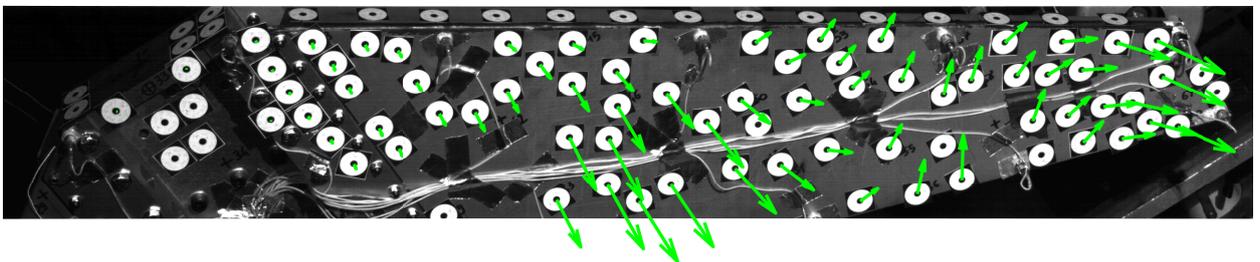


Figure 15: Deflection shape $\underline{\phi}_8^I$

The unknown vector $\underline{q}(t)$ may be computed through the pseudo-inverse of $\underline{\Phi}^{\mathcal{I}}$. Thus $\underline{\Phi}^{\mathcal{I}}$ is required to be full column rank. Although $\underline{\Psi}^{\mathcal{W}}$ in the world frame is an orthogonal basis, its restriction to all the targets, giving $\underline{\Phi}^{\mathcal{W}}$, combined with the perspective projection, has induced the loss of the orthogonality property for $\underline{\Phi}^{\mathcal{I}}$. Moreover, if one keeps too many modal shapes and too few targets, i.e. if $M > 2N$, the full column rank property is certainly lost. Even when $M \leq 2N$, the full column rank is not ensured, depending on the distribution of the targets on the vibrating mechanical structure. For example, the 3D deflections shapes $\underline{\psi}_2^{\mathcal{W}}$ and $\underline{\psi}_3^{\mathcal{W}}$ shown in Figures 7 and 8 are symmetrical modes. As a half of the structure is filmed, the 2D target deflection shapes $\underline{\phi}_2^{\mathcal{I}}$ and $\underline{\phi}_3^{\mathcal{I}}$ shown in Figures 12 and 13 are almost identical. Thus in the present case $\underline{\Phi}^{\mathcal{I}}$ is not full column rank or at least is close to singular. Thus $\underline{\Phi}^{\mathcal{I}}$ must be modified.

For $\underline{\Phi}^{\mathcal{I}}$ to be a basis, we recommend to have $M \ll 2N$ and to remove quasi-colinear columns as described by the algorithm below. This algorithm is inspired by the Gram-Schmidt process, but a vector is kept in the basis only if its angle with the rest of the basis is greater than a given threshold. The major point is to obtain the list $\underline{\ell}$ of vector to be kept in the basis:

1. Initialisation of:

- the index $i = 1$
- the list of vectors to be kept $\underline{\ell} = 1$
- the angle threshold in radian $tol = 5 \frac{\pi}{180}$
- the basis $\underline{T} = \frac{\underline{\phi}_1^{\mathcal{I}}}{\|\underline{\phi}_1^{\mathcal{I}}\|}$.

2. Increment the index $i := i + 1$.

3. Project the i^{th} column on \underline{T} to get:

- its parallel part $\underline{\phi}_{i \parallel}^{\mathcal{I}} = \underline{T} \underline{T}^t \underline{\phi}_i^{\mathcal{I}}$
- its perpendicular part $\underline{\phi}_{i \perp}^{\mathcal{I}} = \underline{\phi}_i^{\mathcal{I}} - \underline{\phi}_{i \parallel}^{\mathcal{I}}$

4. Compute the angle $\alpha = \arctan \left(\frac{\|\underline{\phi}_{i \perp}^{\mathcal{I}}\|}{\|\underline{\phi}_{i \parallel}^{\mathcal{I}}\|} \right)$

5. If $\alpha > tol$:

- Normalize and add its perpendicular part to the basis $\underline{T} := \left[\underline{T}, \frac{\underline{\phi}_{i \perp}^{\mathcal{I}}}{\|\underline{\phi}_{i \perp}^{\mathcal{I}}\|} \right]$
- add the index to the list $\underline{\ell} := [\underline{\ell}, i]$

6. If $i < M$ go back to 2.

$\underline{\ell}$ is the list of the columns $\underline{\Phi}^{\mathcal{I}}$ that produce a basis. Said differently, ℓ is the list of deflection shapes which produce sufficiently different motions that can be distinguished in the image frame. For example, $\underline{\ell}$ contains 2 but not 3, meaning that $\underline{\phi}_2^{\mathcal{I}}$ must be kept but that $\underline{\phi}_3^{\mathcal{I}}$ must be removed. Thus we will now restrict $\underline{\Phi}^{\mathcal{I}}$ to its columns listed in $\underline{\ell}$. Let us denote $\underline{\Theta}^{\mathcal{I}}$ such a matrix and L the length of the vector $\underline{\ell}$. The matrix $\underline{\Theta}^{\mathcal{I}}$ has $2N$ rows and L columns in the image frame. Finally, Equation (14) express the coordinates of all targets in the image frame with the deflections shapes kept in the list ℓ . In the next section, the motion of the mechanical structure observed with only one camera will be reconstructed in the world frame based the kept deflection shapes.

$$\underline{u}^{\mathcal{I}}(t) = \underline{u}_0^{\mathcal{I}} + \underline{\Theta}^{\mathcal{I}} \underline{q}(t) \quad (14)$$

4. Modal amplitude and 3D reconstruction

Let the 2 rows vector $\underline{v}_k^{\mathcal{I}}(t)$ denote the 2D coordinates obtained with the camera for the target k in the image frame, similar to the notation $\underline{u}_k^{\mathcal{I}}(t)$ used for the FE model in Equation (12). Let $\underline{v}^{\mathcal{I}}(t) = [\dots, (\underline{v}_k^{\mathcal{I}}(t))^t, \dots]^t$ denote 2D coordinates of all targets in the image frame \mathcal{I} obtained with the camera. In a least square sense, these coordinates are equal to the ones of the FE model described by Equation (14): $\underline{v}^{\mathcal{I}}(t) = \underline{u}^{\mathcal{I}}(t)$. As a result, the amplitude $\underline{q}(t)$ of deflection shapes is obtained by:

$$\underline{q}(t) = ((\underline{\Theta}^{\mathcal{I}})^t \underline{\Theta}^{\mathcal{I}})^{-1} (\underline{\Theta}^{\mathcal{I}})^t (\underline{v}^{\mathcal{I}}(t) - \underline{u}_0^{\mathcal{I}}) \quad (15)$$

As the extrinsic calibration is performed on the first picture at $t=0$, we have $\underline{u}_0^{\mathcal{I}} = \underline{v}^{\mathcal{I}}(0)$. Thus, Equation (15) can be written:

$$\underline{q}(t) = ((\underline{\Theta}^{\mathcal{I}})^t \underline{\Theta}^{\mathcal{I}})^{-1} (\underline{\Theta}^{\mathcal{I}})^t (\underline{v}^{\mathcal{I}}(t) - \underline{v}^{\mathcal{I}}(0)) \quad (16)$$

If no interpolation had been used on FE mesh to get the position exactly corresponding to the center of the targets, then $\underline{u}_0^{\mathcal{I}}$ would not be exactly equal to $\underline{v}^{\mathcal{I}}(0)$ but just very close to instead. In that case Equation (16) would give better results than Equation (15) since it compensates for the static error occurring when no mesh interpolation are used. However in the present case the authors have used mesh interpolation and both equations give the same results.

Finally, with $\underline{q}(t)$, the 3D coordinates of the whole FE model can be recovered along with time in the world frame \mathcal{W} thanks to Equation (17). This is a modal expansion similar to Equation (8) where the collection of deflection shapes $\underline{\Psi}^{\mathcal{W}}$ is replaced by $\underline{\Psi}_{|\ell}^{\mathcal{W}}$, its restriction to the columns listed in ℓ . Moreover the coefficient s is a scale factor for display purpose:

$$\underline{u}_{\text{EF}}^{\mathcal{W}}(t) = \underline{u}_{\text{EF} 0}^{\mathcal{W}} + s \times \underline{\Psi}_{|\ell}^{\mathcal{W}} \underline{q}(t) \quad (17)$$

In a similar manner, the 2D coordinates of the whole FE model can be recovered along with time in the image frame \mathcal{I} thanks to Equation (18). The collection of deflection shapes

$\underline{\underline{\Psi}}_{|\ell}^{\mathcal{W}}$ is replaced by $\underline{\underline{\Psi}}_{|\ell}^{\mathcal{I}}$, which is its linearized projection in the image frame:

$$\underline{u}_{\text{EF}}^{\mathcal{I}}(t) = \underline{u}_{\text{EF } 0}^{\mathcal{I}} + s \times \underline{\underline{\Psi}}_{|\ell}^{\mathcal{I}} \underline{q}(t) \quad (18)$$

Let us point out that applying the perspective projection, Equations (3) and (4), on $\underline{u}_{\text{EF}}^{\mathcal{W}}(t)$ given by Equation (17), will produce a $\underline{u}_{\text{EF}}^{\mathcal{I}}(t)$ different from the one given by Equation (18). This is due to the linearized deflections shapes $\underline{\underline{\Psi}}_{|\ell}^{\mathcal{I}}$. But since the scale factor s has no physical meaning in Equation (17), the authors preferred to use Equation (18).

5. Results

The choice of the basis $\underline{\underline{\Psi}}^{\mathcal{W}}$ is of great importance. It is worth noticing that it can be a truncated modal basis, or a concatenation of several modal basis obtained for different boundary conditions: free, clamped, ... One can also add some static deflection shapes. In case of concatenation of too close deflection shapes, the selection algorithm presented in Section 3 will avoid any problem. But concatenating too much deflection shapes tends to lose the physical meaning. Here the authors have chosen to concatenate the rigid body translation along the direction of the shaker with the modal basis of the mechanical structure clamped in its center part.

Figure 16 shows the trajectories of each target in the image frame along with time with a scale factor equal to 5. This means that the green lines are obtained by plotting $\underline{v}_k^{\mathcal{I}}(0) + 5 \times (\underline{v}_k^{\mathcal{I}}(t) - \underline{v}_k^{\mathcal{I}}(0))$. It clearly shows that the targets placed at the right end of the wing have the larger displacements due to modal deflections, while the targets placed at the center of the structure just exhibit the rigid motion displacement.

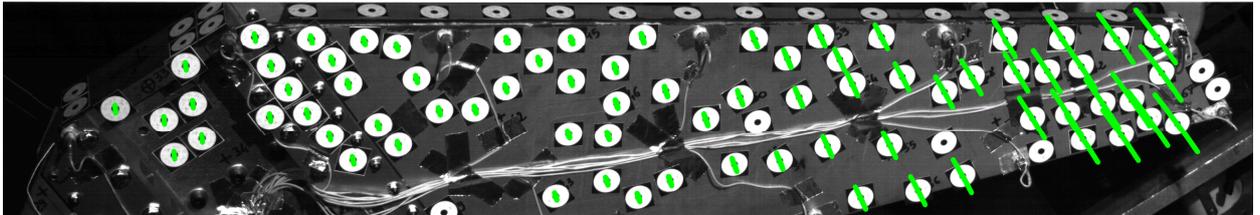


Figure 16: Targets displacement trajectories along with time in the image frame \mathcal{I} , scale factor=5.

Figure 17 shows the evolution along with time of the amplitude $\underline{q}_i(t)$ of some of the retained deflection shapes. The motion seems to be dominated by the second deflection shape, see $\underline{\psi}_2^{\mathcal{W}}$ and $\underline{\phi}_2^{\mathcal{I}}$ in Figures 7 and 12. The rigid body motion is also present with great amplitude around $t = 1.2$ s, due to the shaker motion, see $\underline{\psi}_1^{\mathcal{W}}$ and $\underline{\phi}_1^{\mathcal{I}}$ in Figures 6 and 11. The shaker motion become neglectable around $t = 1.8$ s. The other modes have neglectable contributions to the overall motion.

The method presented here may be subject to some errors due to: mistakes in the intrinsic and extrinsic calibration of the camera, targets detection and tracking, FE model

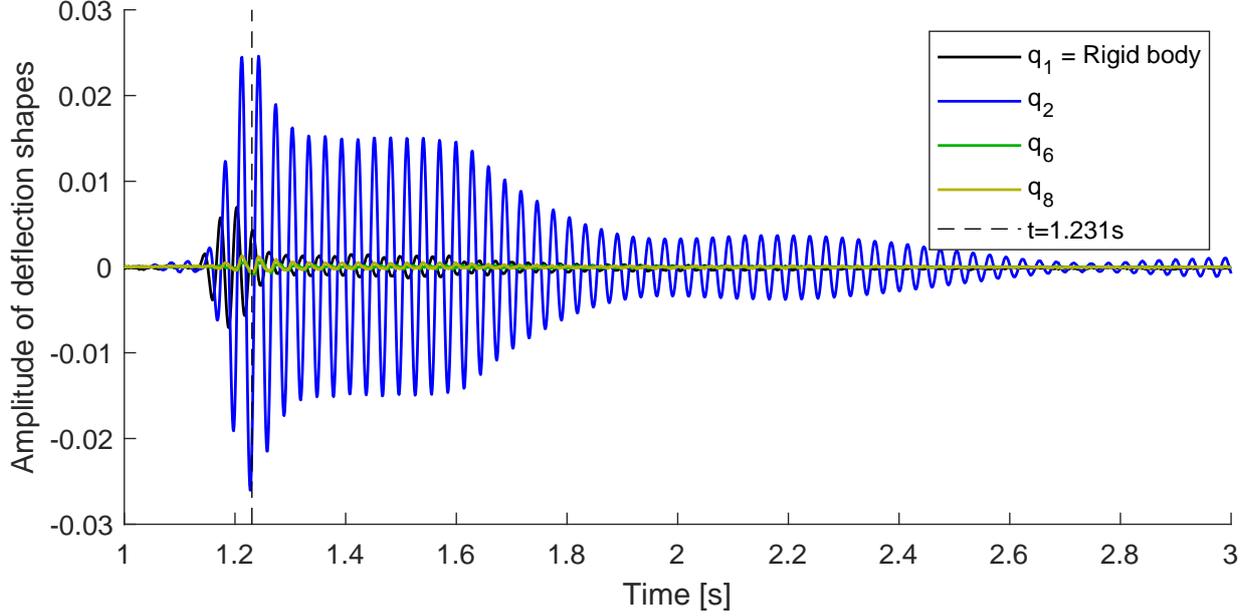


Figure 17: Evolution along with time of the amplitude $\underline{q}_i(t)$ of some of the retained deflection shapes.

accuracy (material, geometry, mesh refinement, boundary conditions, ...), choice of deflection shapes, linearization of the deflection shapes in \mathcal{I} , ... To estimate the displacement error after reconstruction, we can compare the identified motion given by Equation (14) and the motion $\underline{v}^{\mathcal{I}}(t)$ measured with the camera. This comparison is done in the image frame and gives an estimation of the true error occurring in the world frame. The error estimator $\Delta_L(t)$ and the maximal displacement $\Delta_0(t)$ over all targets are given by Equations (19). $\Delta_L(t)$ is obtained when keeping L modes, i.e. $\underline{q}(t)$ has been computed with a L -columns matrix $\underline{\Theta}^{\mathcal{I}}$ in Equation (16).

$$\Delta_L(t) = \max_k \left\| \underline{v}_k^{\mathcal{I}}(t) - \underline{v}_k^{\mathcal{I}}(0) - \underline{\Theta}_{|k}^{\mathcal{I}} \underline{q}(t) \right\| \quad (19a)$$

$$\Delta_0(t) = \max_k \left\| \underline{v}_k^{\mathcal{I}}(t) - \underline{v}_k^{\mathcal{I}}(0) \right\| \quad (19b)$$

Figure 18 shows the evolution of the maximal displacement $\Delta_0(t)$ in blue. Because the structure is vibrating, it goes back to its equilibrium position periodically, thus $\Delta_0(t)$ is equal to zero periodically. This is the reason why a relative error estimator has no sense in this case. The maximal displacement error are plotted in green for $\Delta_2(t)$ and in yellow for $\Delta_{15}(t)$. It can be seen from Figure 18 that $\Delta_2(t)$ and $\Delta_{15}(t)$ are almost equal and very small compared to $\Delta_0(t)$. This means that the two first deflection shapes $\underline{\phi}_1^{\mathcal{I}}$ and $\underline{\phi}_2^{\mathcal{I}}$ are sufficient to describe the motion observed by the camera.

After motion identification with 15 deflection shapes, Figure 19 shows the trajectories of each target residual displacement in the image frame along with time with a scale factor

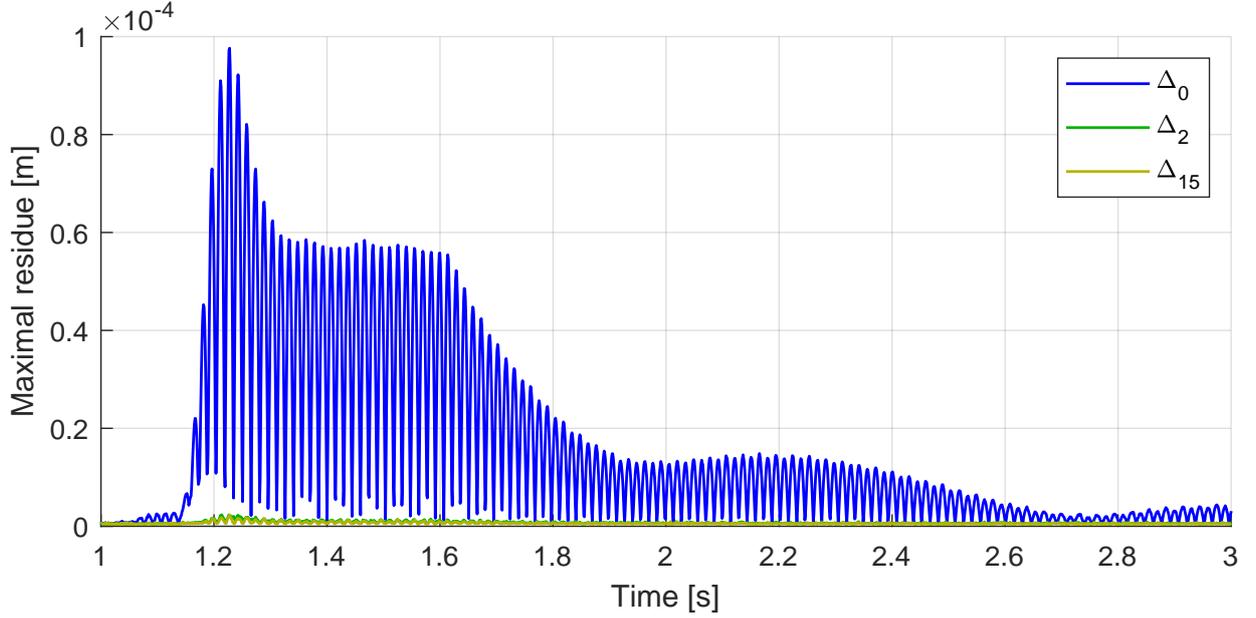


Figure 18: Evolution of the maximal displacement over all targets, Δ_0 in blue and evolution of the maximal displacement errors over all targets, Δ_2 in green and Δ_{15} in yellow when keeping respectively 2 and 15 deflections shapes for the motion identification.

equal to 250. This means that the green lines are obtained by plotting $\underline{v}_k^{\mathcal{I}}(0) + 250 \times (\underline{v}_k^{\mathcal{I}}(t) - \underline{v}_k^{\mathcal{I}}(0) - \underline{\Theta}_{|k}^{\mathcal{I}} \underline{q}(t))$. Let us compare the displacement in Figure 16 with a scale factor equal to 5 and the residual displacement in Figure 19 with a scale factor equal to 250. The biggest green lines have almost the same length, thus the residual displacement is around 2% of the displacement. In Figure 19, along the bolted junction between the center of the structure and the wing, the residual displacements do not seem to be random. The authors think that at least one deflection shape is missing to describe what happens here, maybe local slipping in the bolted junction.

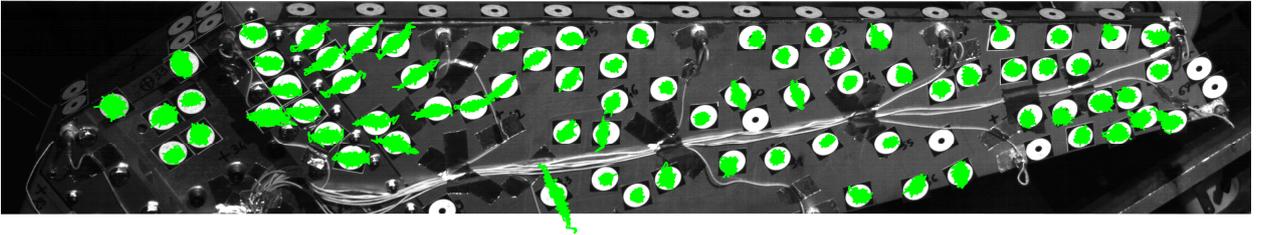


Figure 19: Targets residual displacement trajectories in the image frame \mathcal{I} after motion identification with 15 deflection shapes, scale factor=250.

Figures 20 and 21 shows the 2D and 3D deflection shapes at $t = 1.231$ s reconstructed with Equations (17) and (18) with a scale factor $s = 25$. The time instant $t = 1.231$ s is shown on Figure 17. It corresponds to the largest response of the structure.

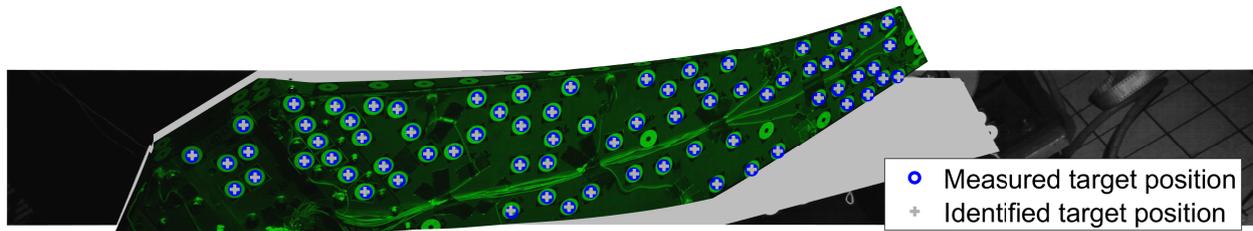


Figure 20: Displacement reconstruction in the image frame \mathcal{I} at $t=1.231$ s, scale factor=25.



Figure 21: Displacement reconstruction in the world frame \mathcal{W} at $t=1.231$ s, scale factor=25.

6. Conclusion

The goal of the work presented in this paper is to reconstruct the 3D motion of a mechanical structure based on the video measurements performed with a single camera and by knowing a collection of deflection shapes of a Finite Element model of the structure. This method can be related to modal analysis.

The intrinsic parameters of the camera have been obtained by following the camera manufacturer procedure. The extrinsic parameters of the camera have been found according to an improvement of the matching point method: by adding matching edges. The motion in the video is measured by tracking targets bounded on the structure with the MSER region algorithm.

The Finite Element deflection shape collection has been transformed from 3D world frame to 2D image frame thanks to linearization of the perspective projection. An algorithm of deflection shape selection has been used to ensure that the collection was a basis, otherwise the method fails. The projection of the targets motion on this basis provides the amplitudes of the deflection shapes along with time. These amplitudes are the same in the 2D image frame

and in the 3D world frame. Finally, the 2D and the 3D motion of the Finite Element model are reconstructed with the selected deflection shapes. Starting from the time evolution of deflection shapes amplitude, one could estimate damping parameters or validate a FE model.

The accuracy of the method is estimated thanks to the residual displacement in the image frame. The residual displacements appear to be small compared to the overall motion. Looking at the residual displacements trajectories can also give some clues to potentially missing deflection shapes.

One major advantage of the method presented here is that it can be applied with several cameras, without any overlapping field of view. Also, this can be used to magnify the vibration amplitudes of a large frequency bandwidth by selecting the corresponding modal basis of the FE model.

7. Acknowledgements

The authors would like to acknowledge Bernard Colomies, Carole Treffot and Floriane Soulas from the SOPEMEA team for their time and expertise. They have been of great help to conduct the vibration tests on their electrodynamic shaker.

References

- [1] F. Renaud, [Graphical abstract of this article](#).
URL <https://www.youtube.com/watch?v=BbI0KsZv0Kw&feature=youtu.be>
- [2] 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 (2015) 58 – 71.
- [3] J. G. Chen, A. Davis, et al, Video camera-based vibration measurement for civil infrastructure applications, *J. Infrastruct. Syst.* 23 (3) (2017) 1–11.
- [4] Y. Yang, C. Dorn, et al, Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification, *Mechanical System and Signal Processing* 85 (2017) 567–590.
- [5] Y. Xu, J. M. W. Brownjohn, Review of machine-vision based methodologies for displacement measurement in civil structures, *Journal of Civil Structural Health Monitoring* 8 (2018) 91–110.
- [6] W. Heidrich, P. Slusallek, H. Seidel, An image-based model for realistic lens systems in interactive computer graphics, 1997, pp. 68–75.
- [7] D. A. Forsyth, J. Ponce, *Computer Vision: A Modern Approach*, Prentice Hall Professional Technical Reference, 2002.
- [8] B. Prescott, G. McLean, Line-based correction of radial lens distortion, *Graphical Models and Image Processing* 59 (1) (1997) 39 – 47.
- [9] F. Devernay, O. Faugeras, Straight lines have to be straight, *Machine Vision and Applications* 13 (2001) 14–24.
- [10] C. Ricolfe-Viala, A.-J. Sánchez-Salmerón, Correcting non-linear lens distortion in cameras without using a model, *Optics & Laser Technology* 42 (2010) 628–639.
- [11] Z. Zangh, A flexible new technique for camera calibration, in: *Microsoft Research*, 1998, pp. 598–71.
- [12] E. Guillou, D. Meneveaux, E. Maisel, K. Bouatouch, Using vanishing points for camera calibration and coarse 3d reconstruction from a single image, *The Visual Computer* 16 (2000) 396–410.
- [13] O. Chenguang, S. S. Shi, Y. Z. You, Z. Kaichun, Extrinsic parameter calibration method for a visual/inertial integrated system with a predefined mechanical interface, *Sensors* 19 (2019) 3086.

- [14] M. Wilczkowiak, E. Boyer, P. Sturm, Camera calibration and 3D reconstruction from single images using parallelepipeds, in: 8th International Conference on Computer Vision (ICCV '01), Jul 2001, pp. 142–148.
- [15] H. Fan, H. Su, L. Guibas, A point set generation network for 3d object reconstruction from a single image, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2463–2471.
- [16] T. Bebernis, D. Ehrhardt, High-speed 3d digital image correlation vibration measurement: Recent advancements and noted limitations, *Mechanical Systems and Signal Processing* 86 (B) (2016) 35–48.
- [17] W. Wang, et al., Frequency response functions of shape features from full-field vibration measurements using digital image correlation, *Mechanical Systems and Signal Processing* 28 (2012) 333–347.
- [18] P. Reu, D. Rohe, L. Jacobs, Comparison of dic and ldv for practical vibration and modal measurements, *Mechanical Systems and Signal Processing* 86 (2017) 2–16.
- [19] B. Pan, L. Yu, Q. Zhang, Review of single-camera stereo-digital image correlation techniques for full-field 3d shape and deformation measurement, *Sci. China Technol. Sci.* 61 (2018) 2–20.
- [20] L. Yu, B. Pan, Single-camera high-speed stereo-digital image correlation for full-field vibration measurement, *Mechanical Systems and Signal Processing* 94 (2017) 374 – 383.
- [21] T. Durand-Texte, M. Melon, E. Simonetto, S. Durand, M.-H. Moulet, Single-camera single-axis vision method applied to measure vibrations, *Journal of Sound and Vibration* 465 (2020) 115012.
- [22] H. Festjens, G. Chevallier, J.-L. Dion, A numerical tool for the design of assembled structures under dynamic loads, *International Journal of Mechanical Sciences* 75 (2013) 170 – 177.
- [23] H. Festjens, G. Chevallier, J.-L. Dion, Nonlinear model order reduction of jointed structures for dynamic analysis, *Journal of Sound and Vibration* 333 (7) (2014) 2100 – 2113.
- [24] G. Chevallier, H. Festjens, J.-L. Dion, N. Peyret, Model reduction and lumped models for jointed structures, in: M. Allen, R. Mayes, D. Rixen (Eds.), *Dynamics of Coupled Structures, Volume 1*, Springer International Publishing, Cham, 2014, pp. 273–279.
- [25] N. Peyret, J.-L. Dion, G. Chevallier, A framework for backbone experimental tracking : Piezoelectric actuators, stop-sine signal and kalman filtering, *Mechanical Systems and Signal Processing* 78 (2016) 28 – 42, special Issue on Piezoelectric Technologies.
- [26] J. Matas, O. Chum, M. Urban, T. Pajdla, Robust wide-baseline stereo from maximally stable extremal regions, *Image and Vision Computing* 22 (10) (2004) 761 – 767, british Machine Vision Computing 2002.
- [27] HALCON, Solution Guide III-C - 3D Vision, MVTEC.
- [28] Z. Shang, Z. Shen, Multi-point vibration measurement and mode magnification of civil structures using video-based motion processing, *Automation in Construction* 93 (2018) 231 – 240.
- [29] N. Wadhwa, H.-Y. Wu, A. Davis, et al., Eulerian video magnification and analysis, *Commun. ACM* 60 (1) (2016) 87–95.
- [30] J. C. K. Chou, Quaternion kinematic and dynamic differential equations, *IEEE Transactions on robotics and automation* 8 (1) (1992) 53–64.
- [31] A. Goeller, Contribution à la perception augmentée de scènes dynamiques : schémas temps réels d’assimilation de données pour la mécanique du solide et des structures, Ph.D. thesis (2018).