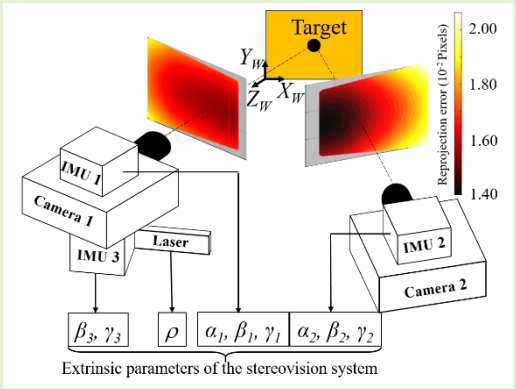
Sensor-Based Calibration of Camera's Extrinsic Parameters for Stereophotogrammetry

Fabio Bottalico, Christopher Niezrecki, Kshitij Jerath, *Member, IEEE*, Yan Luo, *Member, IEEE*, and Alessandro Sabato, *Member, IEEE*

***Abstract*—Stereophotogrammetry is a well-recognized technique for structural health monitoring. Before performing any quantitative stereophotogrammetry measurement, the cameras must be calibrated to obtain the intrinsic and extrinsic parameters of the stereovision system. However, when large-sized structures are to be monitored, the calibration process is challenging and limits the use of stereophotogrammetry. In this research, a sensor-based calibration method for determining the extrinsic parameters of a stereovision system is presented and validated. A multi-sensor board has been developed that synchronizes inertial measurement units (IMUs) and a laser on a single board computer to measure the spatial orientation and the distance of two paired cameras and compute the extrinsic parameters of the stereovision system. The effectiveness of the sensor-based calibration is evaluated through both analytical studies to quantify the effects of performance degradation caused by the sensors' noise as well as laboratory tests. Results show that the sensor-based calibration is effective in quantifying displacement with errors below 3% when compared to measurements performed using a stereovision system calibrated with the traditional image-based procedure.**



***Index Terms*— Camera calibration, digital image correlation, inertial measurement unit, laser, stereophotogrammetry.**

I. Introduction

STEREOPHOTOGRAMMETRY has become a popular tool for non-destructive inspection and structural health monitoring thanks to progress made in image-processing algorithms [1, 2] and advantages that this approach provides to overcome some of the limitations of traditional contact-based methods (e.g., discrete and limited number of measurement points, sensor power requirements and data transmission, installation challenges, and durability) [3, 4]. To perform stereophotogrammetry measurements using techniques such as three-dimensional digital image correlation (3D-DIC), a set of image pairs is taken during an experiment with two synchronized cameras. The first pair of images is used as the reference to which all subsequent photos are compared [5]. By tracking the pixel motion of each feature identified in the image set, the in-plane and out-of-plane strain as well as the shape, deformation, and displacement of the targeted structure can be measured [6]. Over the years, stereophotogrammetry has been used successfully to monitor several types of structures in the civil [7-9], transportation [10, 11], mechanical [12], and energy [13] engineering domains.

Before performing stereophotogrammetry, the stereovision system needs to be calibrated to obtain the relative spatial position (i.e., extrinsic parameters) of the two cameras [14, 15] and account for the internal distortion of each lens (i.e., intrinsic parameters) [16]. For this reason, a sequence of images of a calibration object containing optical targets or a pattern whose features or locations are well-known is captured with the cameras at different positions and orientations [17] prior to performing calibration using a method described by Zhang [18]. By identifying the location of corresponding points from the calibration object in both images, the extrinsic and intrinsic parameters are extracted using the bundle adjustment [19]. Some examples for determining the intrinsic and extrinsic parameters include the use of 3D calibration objects such as multiplanar targets [20], cylinders [21], scale bars [22], active phase targets [23], single feature points [24, 25]. While those approaches proved to be effective at calibrating a stereovision system, once the last image of the calibration object is recorded, the cameras' relative position cannot change, otherwise a loss of calibration will occur. For this reason, cameras are mounted to a rigid bar or a set of fixed camera mounts to prevent any relative motion. Furthermore, to generate a calibrated measurement volume, the dimensions of the calibration object need to be comparable with the size of the system being tested.

This paragraph of the first footnote will contain the date on which you submitted your paper for review. (Corresponding author: Alessandro Sabato.)

F. Bottalico, C. Niezrecki, K. Jerath, and A. Sabato are with the Department of Mechanical Engineering at the University of Massachusetts Lowell, Lowell, MA 01852 USA (email: Fabio_Bottalico@student.uml.edu; Christopher_Niezrecki@uml.edu; Kshitij_Jerath@uml.edu; Alessandro_Sabato@uml.edu). Y. Luo is with the Department of Electrical and Computer Engineering at the University of Massachusetts Lowell, Lowell, MA 01852 USA (email: Yan_Luo@uml.edu).

For example, for a field of view (FOV) larger than 3 m, a large-area calibration procedure is performed that requires building a customized calibration object which can make the process lengthy, complicated, and requires multiple person-hours of work [26]. Alternatively, the cameras can also be calibrated using a set of scale bars on the object of interest, however if the object is large, the scale bars can be difficult to measure. Regardless, for both approaches, if the cameras are moved, the calibration will contain gross spatial measurement errors.

The difficulties involved in the traditional image-based calibration summarized above make quantitative stereophotogrammetry measurements on large-sized structures difficult or impractical and provide the motivation for this research. In this study, a multi-sensor system with three inertial measurement units (IMUs) and a laser sensor has been developed and installed on the cameras of a stereovision system to determine their extrinsic parameters. The accuracy of the proposed system is validated analytically to quantify the effects of performance degradation caused by the IMU and laser sensors' noise. In addition, the efficacy of the proposed system in performing 3D-DIC measurements is compared with the results obtained from a stereovision system calibrated using a traditional image-based procedure.

In this paper, the mathematical foundation of stereophotogrammetry calibration and how the framework is extended for sensor-based calibration are reviewed within Section II. Section III describes the experiments performed to characterize the noise floor of the IMU and laser sensors used to extract the extrinsic parameters. The results of the analytical simulation used to quantify the effects of performance degradation caused by the sensors' noise are summarized in Section IV. Section V presents the back-to-back comparison between a measurement performed using the sensor-based calibration and a 3D-DIC measurement performed using traditional image-based calibration. Finally, with Section VI future work and conclusions are presented. To the authors' best knowledge, the work presented in this paper is the first successful attempt at computing the full set of extrinsic parameters of a stereovision system using only data extracted from sensors, without having to rely on images.

II. Principles of Stereophotogrammetry Calibration

Photogrammetry is based on the pinhole camera model that relies on the projection of a 3D point P with coordinates $\boldsymbol{P} = (P_x,\ P_y,\ P_z)^{\mathrm{T}}{}_W$ into a 2D point $p = (p_u,\ p_v)^{\mathrm{T}}$ in the pixel coordinates on the camera's retinal plane [27, 28]. This 3D-to-2D transformation requires a knowledge of the camera intrinsic and extrinsic parameters as expressed in:

$$\begin{bmatrix} p_u \\ p_v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & \theta & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_{xx} & R_{xy} & R_{xz} & T_x \\ R_{yx} & R_{yy} & R_{yz} & T_y \\ R_{zx} & R_{zy} & R_{zz} & T_z \end{bmatrix}_W \begin{bmatrix} P_x \\ P_y \\ P_z \\ 1 \end{bmatrix}_W \tag{1}$$

where p_u and p_v are the pixel coordinates of point p, f_x and f_y are intrinsic parameters such that $f_x = F \cdot S_x$ and $f_y = F \cdot S_y$, where $\boldsymbol{F}$ is the camera's focal length in mm and S_x and S_y are the sensors' scale factors in pixels/mm in the x and y directions, respectively. In Eq. 1, c_x and c_y are the optical center of the image sensor, θ the skew factor of the image, and R_{ij} and T_i the components of the rotation matrix **R** and the translation vector **T** in the global frame of reference W [16]. While **R** and **T** define the extrinsic parameters of the camera, f_x, f_y, c_x, c_y, and θ define the camera's intrinsic parameters.

When a stereovision system is considered, two pinhole cameras look at the same scene from slightly different perspectives. The 3D point P is projected to two 2D points p_i, one for each camera's retinal plane [i.e., $p_1 = (p_{u1}, p_{v1})$ and $p_2 = (p_{u2}, p_{v2})$]. To perform the 3D reconstruction of a point P in W, Eq. 2 can be used when the projections of point P in the retinal planes p_1 and p_2 are known:

$$P_{Rc}^{W} = \boldsymbol{R}_1^W\, P_{Rc}^1 = \boldsymbol{R}_1^W [(\boldsymbol{M}^T \boldsymbol{M})^{-1}\, \boldsymbol{M}^T\, \boldsymbol{b}] \tag{2}$$

where P_{Rc}^{W} is the reconstructed point in space in the global frame of reference W, $\boldsymbol{R}_1^W$ is the rotation matrix of camera 1 in the global frame of reference, and P_{Rc}^1 is the reconstructed point in space in the frame of reference of camera 1. As expressed in Eq. 2, P_{Rc}^1 can be calculated by determining **M** and **b** using a linear least squares method. Both **M** and **b** are function of the intrisic and extrinsic parameters as defined by equation (9.2) in [29].

Because the coordinates of point P are triangulated using its projections p_1 and p_2, it is fundamental to know the relative position, orientation, and settings of the two cameras by determining their extrinsic and intrinsic parameters. For a stereovision system, intrinsic and extrinsic parameters can be decoupled and computed indipendently. In particular, the measurement of the intrinsic parameters for two cameras focused at the hyper-focal distance is a one-time operation and its complexity is independent of the FOV of the targeted object. For this reason, this research focuses on developing a more time-efficient way to estimate the extrinsic parameters and potentially saving hours of work when stereophotogrammetry is performed on large-scale structures (e.g, greater than ~3 m).

To determine **R** and **T**, traditional calibration requires taking pictures of a calibration object. Because the complexity of this operation is function of the FOV of the targeted object, novel approaches that rely on sensors have been proposed. However, those approaches still require pictures to estimate some of the extrinsic parameters [30] or evaluate only a subset of them without requiring pictures [31, 32]. A more recent solution involves creating a virtual calibration target having the size of the desired FOV after measuring a subset of the extrinsic parameters with two IMUs and three laser sensors attached to the cameras [33]. Using a virtual calibration target still requires a new calibration every time cameras are moved. In addition, the approach described in [33] constrains the cameras' baseline to be parallel to the ground, limiting the flexibility of the method. The solution presented in this paper relies on a multi-sensor system which includes three IMUs and a laser sensor embedded on two RaspberryPi 4 installed on the cameras of the stereovision system (see **Figure 1a**). The core components of the multi-sensor board include:

- an ultra-low noise 9-axis LPMS-IG1 IMU by LP-Research that includes a 3-axis gyroscope, a 3-axis

accelerometer, and a 3-axis magnetometer with the possibility to **i)** use a Madgwick [34]or a Kalman filter, **ii)** activate/deactivate the magnetometer, and **iii)** activate/deactivate the gyroscope's auto-calibration functionality to remove the yaw's drift [35];
- a M88B laser module by JRT Meter Technology [36];
- a 2 Megapixel Basler puA1600-60uc camera with pixel size of 4.5x4.5 µm [37];
- a Master Station that takes care of the sensor synchronization and data acquisition.

As shown in **Figure 1b**, IMU 1 and IMU 2 are used to measure the cameras' orientation (i.e., roll α_i, pitch β_i, and yaw γ_i of camera 1 and camera 2), while IMU 3 and the laser evaluate the distance between the cameras. The components of the rotation matrix $\boldsymbol{R}_i^W$ of the i^{th} camera can be calculated in the global frame of reference W from the angles α_i, β_i, and γ_i measured with IMU 1 and IMU 2 using:

$$\boldsymbol{R}_i^W = \boldsymbol{R}_z(\gamma_i)\,\boldsymbol{R}_y(\beta_i)\,\boldsymbol{R}_x(\alpha_i) = \begin{bmatrix} cos(\gamma_i) & -sin(\gamma_i) & 0 \\ sin(\gamma_i) & cos(\gamma_i) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} cos(\beta_i) & 0 & sin(\beta_i) \\ 0 & 1 & 0 \\ -sin(\beta_i) & 0 & cos(\beta_i) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\alpha_i) & -sin(\alpha_i) \\ 0 & sin(\alpha_i) & cos(\alpha_i) \end{bmatrix} \tag{3}$$

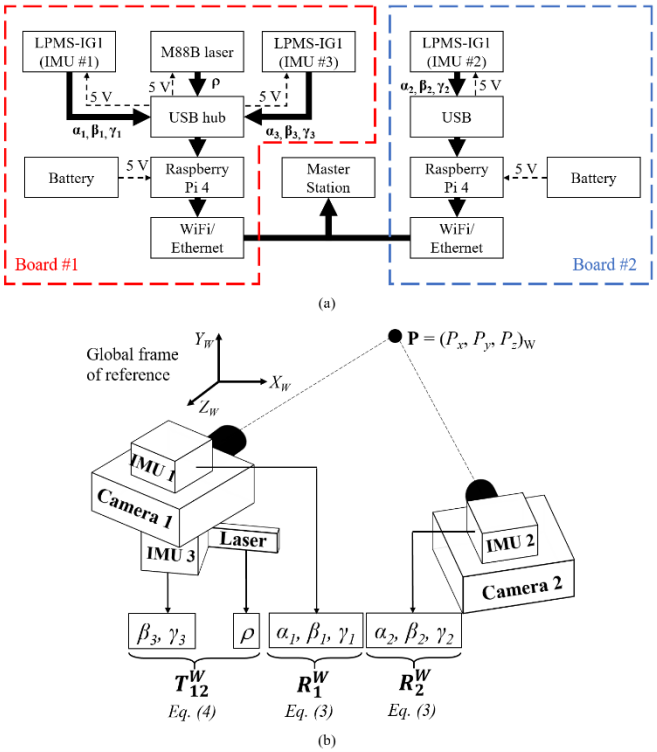


Fig. 1. Proposed multi-sensor board: a) block-diagram showing components and signal flow and b) schematic representation of the developed sensor-based calibration method used to measure the extrinsic parameters.

By using IMU 3 and the laser, the translation vector $\boldsymbol{T}_{12}^W$ is determined in spherical coordinates $[\rho, \gamma_3, \beta_3]^T$, where ρ is the distance measured with the laser, γ_3 is the laser's azimuth and β_3 is the lasers' elevation. As a result, $\boldsymbol{T}_{12}^W$ can then be expressed in Cartesian coordinates using:

$$\boldsymbol{T}_{12}^W = \begin{bmatrix} \rho\, sin(\gamma_3)\, cos(\beta_3) \\ \rho\, sin(\gamma_3)\, cos(\beta_3) \\ \rho\, cos(\gamma_3) \end{bmatrix} \tag{4}$$

The use of IMU 3 and the laser allows performing stereophotogrammetry without any limitations on the position of camera 2 in the 3D space, thus removing the constraint of requiring a rigid connection between the two cameras. As opposed to the approach presented in [33], in the current approach the baseline of the two cameras does not need to be parallel to the ground. Because all IMUs evaluate the Euler angles in the global frame of reference W, $\boldsymbol{T}_{12}^W$ must be expressed in camera 1's frame of reference as $\boldsymbol{T}_{12}^1$ to be used for stereophotogrammetry. This can be done with:

$$\begin{cases} \boldsymbol{T}_{12}^1 = (\boldsymbol{R}_1^W)^T \ \boldsymbol{T}_{12}^W \\ \boldsymbol{R}_2^1 = (\boldsymbol{R}_1^W)^T \boldsymbol{R}_2^W \end{cases} \tag{5}$$

where $(\boldsymbol{R}_1^W)^T$ is the transpose of the rotation matrix that identifies the orientation of camera 1 in W (i.e., $\boldsymbol{R}_1^W$) computed from the IMU 1 data using Eq. 3, $\boldsymbol{R}_2^W$ the rotation matrix calculated from the data measured by IMU 2 in W, and $\boldsymbol{R}_2^1$ the rotation matrix of camera 2 in the frame of reference of camera 1. It should be noted that Eq. 5 describes a case in which the coordinate systems of all sensors attached to camera 1 are centered in the camera's frame of reference. The decision to position the origin of camera 1 in the origin of W and the origin of the laser in the origin of camera 1 is to simplify the computational load. In contrast, the coordinate system of IMU 2 corresponds to the frame of reference of camera 2. When working with actual sensors with non-negligible physical sizes, the relative translations and rotations between the different sensors can be calculated using a chain of homogeneous transformations computed with the Denavit-Hartenberg (DH) parameters [38]: as the camera 1-to-laser assembly includes a pan-tilt mechanism and is not unlike a robotic arm, the use of equations pertaining to robotics is justified. Once the extrinsic parameters are measured, Eq. 2 is used to reconstruct the 3D position of point P given its projections in the retinal plane of the two cameras.

III. Experimental Characterization of The Multi-sensor Board

By using the setup shown in **Figure 1**, it is possible to extract the extrinsic parameters of the stereovision system. Because those parameters are affected by the sensors' noise, laboratory experiments were performed to quantify the effect of the IMUs and laser's noise floor on the accuracy of the calibration. **Figure 2** shows the setups used for characterizing the sensors while the results of the tests are discussed in this section.

A. IMU characterization

The experiments performed on the LPMS-IG1 consisted of static measurements to determine the noise floors of the IMU at rest when different configurations (i.e., combinations of enabled/disabled magnetometer and gyroscope) are used. During the tests, the IMU was mounted on a heavy block (see **Figure 2a**) to minimize the effect of ambient vibrations, data was collected for five minutes at a sampling rate of 100 Hz, and

the test was performed 12 times for repeatability. The five-minute interval was chosen as a reasonable period for allowing statistical analysis of the sampled data, considering that a significantly shorter time interval (e.g., 2 seconds) is enough to extract the extrinsic parameters of the stereovision system for calibration. Eight IMU configurations were tested, but only two are shown in this paper for the sake of brevity.

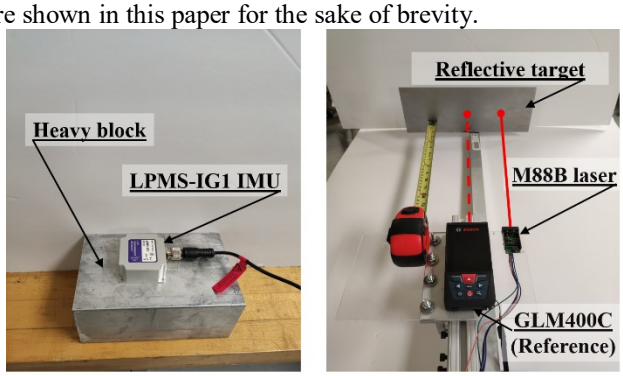


Fig. 2. Experimental setup for the characterization tests of the multi-sensor board: a) IMU setup to identify the noise floor and b) laser setup comparing to a reference laser measure and tape measure.

In particular, Configuration #1 had the auto-calibration enabled and magnetometer disabled, Configuration #2 had the auto-calibration and magnetometer both enabled. A complete description of the eight configurations is shown in [39]. The comparison between the two selected configurations in terms of signal standard deviation (σ) is shown in **Figure 3**, which plots the angular outputs of the IMU over time.

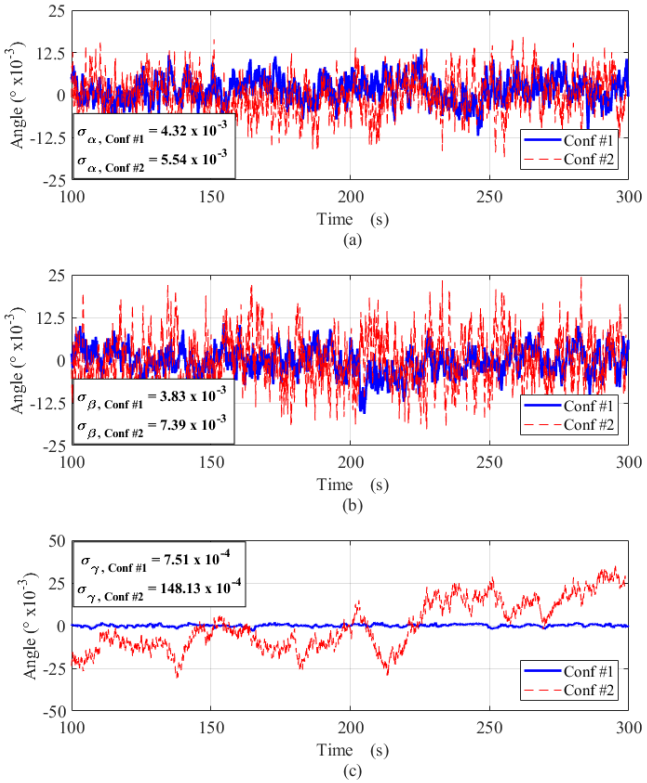


Fig. 3: Comparison of two different IMU's settings on the accuracy of the estimated angles: a) roll, b) pitch, and c) yaw, for Configurations 1 and 2.

It should be noted that the first 100 seconds of the measurement have been omitted for a better comparison between the two selected configurations, as the magnetometer used in Configuration #2 takes time to orient around the East-North-Up (ENU) frame of reference. The results in **Figure 3** show how the magnetometer introduces extra noise in the yaw angle estimation most likely due to electromagnetic interference and orthogonality errors caused by the assembly of the sensor. In particular, the auto-calibration allows for a drift of 4.5×10^{-4} deg over five minutes, as opposed to a value of $\sim 5 \times 10^{-2}$ deg when the magnetometer is enabled. Because of the superior performance of the IMU when no magnetometer is used, the noise floor ranges determined for Configuration #1 that represents the precision of the measurement (i.e., $\sigma_\alpha = \pm 4.32 \times 10^{-3}$ deg.; $\sigma_\beta = \pm 3.83 \times 10^{-3}$ deg.; and $\sigma_\gamma = \pm 0.75 \times 10^{-3}$ deg.), will be used as error ranges to analyze how the accuracy of the calibration is affected by the IMU noise.

B. Laser characterization

A comparison was performed with a commercially available laser measure GLM400C by Bosch to quantify the accuracy of the M88B laser sensor embedded on the multi-sensor board. The two devices were mounted on a support placed on a slider rail pointing perpendicular towards a reflective target to make sure both sensors were able to detect the retro reflected spot (see **Figure 2b**). During the test, the support with the two laser sensors was moved away from the reflective target at ~0.5 m increments. For each distance, twenty-five measures were taken with each sensor. The data collected with the M88B were averaged at each distance and used to determine the noise floor of the measurement by computing the average standard deviation between the distance calculated using the M88B laser ant the regression line of the points measured using the reference GLM400C (see **Figure 4**).

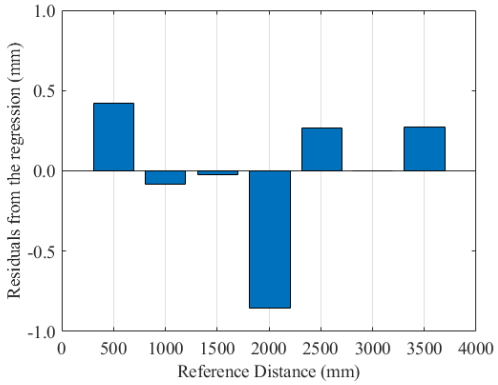


Fig. 4: Residuals of the M88B laser from the regression line calculated from the distances measured using the reference GLM400C.

From the data shown in Figure 4, it is estimated that the noise floor of the M88B laser, σ_{laser} is equal to 4.6×10^{-4} m when compared to the GLM400C laser used as a reference. For this case, the standard deviation was calculated using the *[1/(N-1)]* definition and its value modified using the unbiased estimator correction factor to account for a sample size of N=6. Similarly to what is done for the results obtained for the IMU, σ_{laser} will be used as the error range to analyze how the accuracy of the calibration is affected by inaccuracies in the laser measurements.

IV. Analytical Evaluation of Sensors' Noise on Calibration Accuracy

An analytical study was performed to quantify the effect of the sensors' noise floor on the extrinsic parameters of a stereovision system. The accuracy of the sensor-based calibration is evaluated in terms of reconstruction error ε_{Rc} and reprojection error ε_{Rp}. **Figure 5** provides a graphical representation of ε_{Rc} and ε_{Rp} for easier visualization. According to the literature, ε_{Rp} below 0.5 pixels [40-43] and ε_{Rc} equal to ~1/100 pixel in the in-plane direction and ~1/30 pixel in the out-of-plane direction [44] are acceptable for most stereophotogrammetry measurements. As discussed in Section II, when a point P with known 3D coordinates is projected onto the cameras' retinal planes, two points p_1 and p_2 are generated.

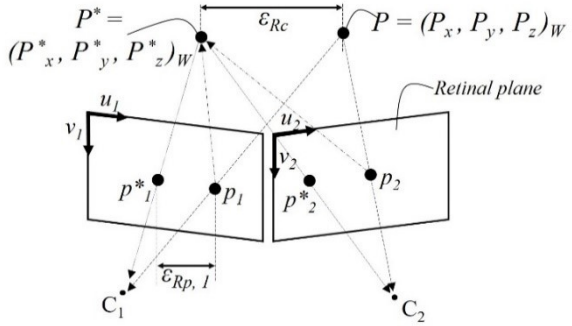


Fig. 5: Representation of the reconstruction and reprojection errors caused by the IMU and laser's noise affecting the extrinsic parameters.

The pixel coordinates (p_{u1}, p_{v1}) and (p_{u2}, p_{v2}) of those points, combined with the orientation and position of the two cameras obtained from the sensors' data are used to reconstruct the 3D coordinates of point P^* using Eq. 2. Because of errors in the reconstruction of the cameras' orientation and positions caused by noise in the IMU and laser as well as lens distortion, point P^* will be offset from the true point P by a factor equal to ε_{Rc}. Once P^* is calculated, it is reprojected onto the cameras' retinal plane using the same estimated extrinsic parameters. This reprojection yields points p_1^* and p_2^*, which can be compared with the corresponding true values p_1 and p_2. The Euclidean distance in pixel between these two sets of data is ε_{Rp} [45].

A. Details of the analytical study

To evaluate the feasibility of the proposed sensor-based calibration and understand the effect of sensor's noise on performance degradation, a stereovision system having a pair of 2 Megapixel Basler puA1600-60uc cameras fitted with 12.5 mm focal length lenses was considered for the analytical simulation. As shown in **Figure 6**, the cameras were set to have a separation angle $\gamma_1 + \gamma_2$ of 25°, base distance ρ equal to 0.75 m, and a working distance of 1.8 m. The cameras were assumed to be planar, thus with α and β equal to 0° (i.e., no roll or pitch). The simulated stereovision system results in a FOV of 0.6 x 0.8 m, where each pixel on the image corresponds to ~1.4 x10^{-3} m in the global frame of reference. Given the specification of the cameras and the obtained FOV, a theoretical in-plane accuracy of ~14 x10^{-6} m and an out-of-plane accuracy of ~46 x10^{-6} m is expected. Those values are used as a reference for ε_{Rc} and ε_{Rp} when the extrinsic parameters are calculated using the sensors' data.

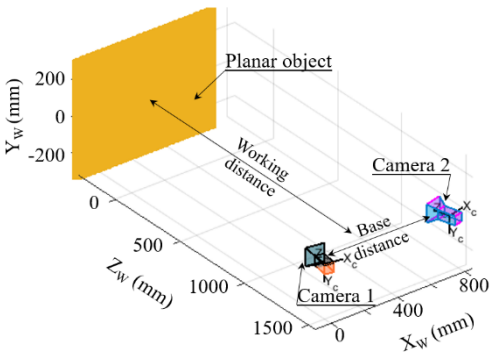


Fig. 6: Stereovision system setup pointing at a planar object used to simulate a 3D-DIC measurement and validate the effect of sensor's noise on the accuracy of the calibration.

In the simulation, the **i)** intrinsic parameters of the stereovision system and **ii)** the coordinates of the points P_i on the surface of the targeted planar object were assumed to be perfectly known and with zero variance noise. The extrinsic parameters were considered to be affected by the random noise of the IMU and laser sensors used to measure the position and orientation of the stereovision system. To generate a series of camera parameters and reconstruct the coordinates of points P_i^*, 100 readings from the IMUs and ten from the laser were generated by randomly varying α, β, γ and ρ in the noise range $\pm\sigma_i$ calculated in Section III (i.e., σ_α, σ_β, σ_γ, σ_{laser}). To provide a statistically significant sample, this operation was repeated 100 times. A total of 10,000 IMU and 1,000 laser values were then used in Eqs. 3 and 4 to compute the extrinsic parameters affected by noise and obtain P_i^* by solving Eq. 2. **Table I** summarizes the orientations and positions of cameras 1 and 2 used in the analytical study and the ranges in which the extrinsic parameters were varied, while **Figure 7** shows the workflow used to compute ε_{Rc} and ε_{Rp}.

TABLE I
Reference Values of Camera Orientation and Position and Noise Ranges Used In the Analytical Study

		α (°)	β (°)	γ (°)	ρ (m)
Reference	Camera 1	0.0	0.0	-10.0	0.0
	Camera 2	0.0	0.0	15.0	0.75
	Laser's IMU	0.0	0.0	-90.0	-
Noise	$\pm\sigma_i$ (x10^{-3})	4.3	3.8	0.7	0.5

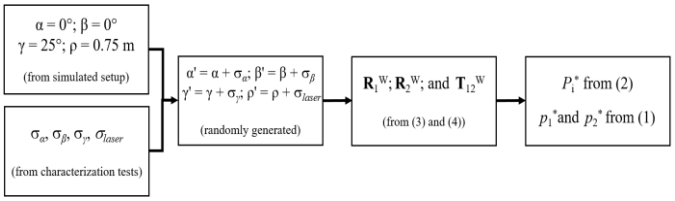


Fig. 7: Workflow of the analytical study used for computing ε_{Rc} and ε_{Rp}.

B. Results of the analytical study

Once the position and orientation of the cameras affected by the sensors' noise is calculated, by using Eq. 1 it is possible to reproject the coordinates of points P_i^* onto the retinal planes to obtain p_i^*. The result of the reprojection is shown in **Figure 8**, which represents the reprojection error maps for the two images recorded with camera 1 and camera 2.

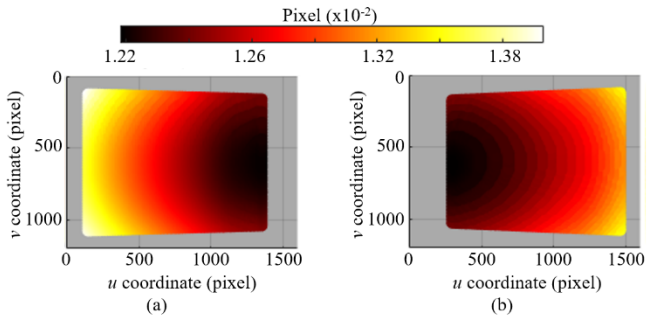


Fig. 8: Reprojection error maps on the retinal plane showing errors on the order of 10^{-2} pixels: a) camera 1 and b) camera 2.

From **Figure 8**, it is observed that ε_{Rp} is on average equal to $1.3\text{x}10^{-2}$ pixel and always below the $4\text{x}10^{-2}$ pixel threshold required for stereophotogrammetry. In addition, the mean in-plane accuracy is equal to $3\text{x}10^{-6}$ m, while the out-of-plane accuracy is equal to $16\text{x}10^{-6}$ m. It should be noted that because $\boldsymbol{R}_i^W$ is affected by the noise of IMU 1, ε_{Rp} of camera 1 is not zero.

To complete the characterization of the sensor-based calibration approach, a sensitivity analysis was performed by considering the effects of increasing noise levels in the IMUs and laser. For this reason, an analytical simulation was run by considering sensors' noise level varying in the $1\sigma_i$, $2\sigma_i$, and $3\sigma_i$ ranges. The reconstruction and reprojection errors were evaluated by keeping constant the noise range of one of the two sensors and increasing the noise range of the other. For example, during the first simulation, the noise floor of the IMU was kept constant to 1σ, while the noise floor of the laser was increased from 1 to 3σ. Then the noise level for the IMU was increased to 2σ and the operation was repeated. The goal of this simulation was to observe how different levels of sensors' noise affect ε_{Rp} and ε_{Rc} and how those values compare to the thresholds available the literature [39, 43]. **Figure 9** provides a summary of the sensitivity analysis compared to the theoretical thresholds for accurate measurements on a 0.6x0.8 m FOV (i.e., dotted lines). In the figure, noise floors generating values to the left of or below the dotted lines have reprojection and reconstruction errors smaller than the threshold.

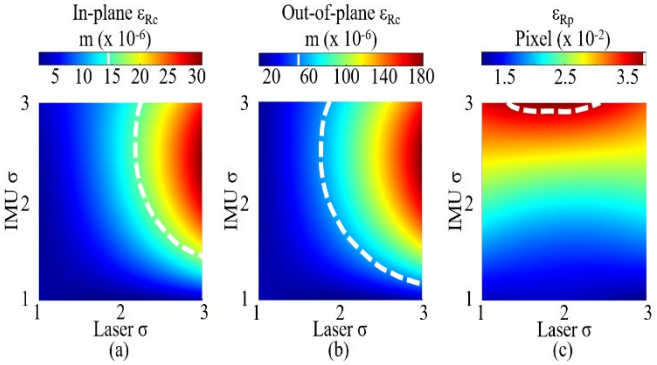


Fig. 9: Results of the sensitivity analysis showing that the multi-sensor system provides sufficient accuracy for stereophotogrammetry for an IMU noise floor up to 3σ and laser noise floor up to 2σ: a) in-plane reconstruction error, b) out-of-plane reconstruction error, and c) reprojection error.

From **Figure 9a** and **Figure 9b**, it can be observed that ε_{Rc} is more sensitive to degradation in the laser data than in the IMU. However, for the majority of the cases simulated, ε_{Rc} is below the thresholds required for accurate measurements. Only when the laser noise floor is above 2σ, then ε_{Rc} is higher than the theoretical accuracy limits. For what concerns ε_{Rp}, this parameter is always below the 0.5-pixel limit and only when the IMU noise is in the 3σ range the error is above $4\text{x}10^{-2}$ pixel (see **Figure 9c**).

V. Experimental Validation of the Proposed Sensor-Based Calibration

A laboratory test was performed to validate the accuracy of the sensor-based calibration approach for 3D-DIC when the extrinsic parameters of a stereovision systems are extracted using the proposed multi-sensor board. **Figure 10a** shows a prototype of the proposed system that consists of two 2 Megapixel Basler puA1600-60uc cameras fitted with 12 mm lenses, three LPMS IG1 IMUs, one M88B laser distance sensor, and two Raspberry Pi 4. The Raspberry Pi's are used for data synchronization and acquisition. Because the settings selected for the IMUs disable the magnetometers (see Section III.A), the sensors do not measure angles in the ENU global coordinate system. Instead, the coordinate system depends on the initial position of each individual IMU. For this reason, before measuring the extrinsic parameters, all IMUs must be initialized to a user-defined global coordinate system using the calibration zeroing block shown in **Figure 10a**. The block provides a perpendicular square edge that enables the pitch and yaw signals of the IMUs to be zeroed prior to measurement and thereby share the same coordinate system when moved.

To validate the accuracy of the multi-sensor system, a 3D-DIC measurement was performed using the setup shown in **Figure 10b**. The experiment replicates the setup shown in **Figure 6** used for the analytical study with two cameras placed in front of a ~0.6 x 0.8 m planar object, which was moved in the negative x-direction from zero to $1\text{x}10^{-2}$ m at increments of $0.5\text{x}10^{-3}$ m. The planar object is speckled with a random black-and-white pattern to allow correlating the left and right images and performing 3D-DIC. During the test, five images were acquired with both cameras for each displacement location of the planar object to allow for data averaging.

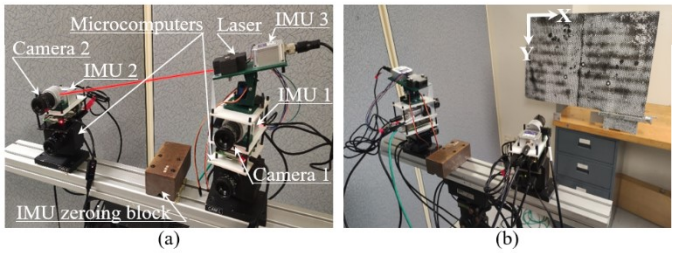


Fig. 10: Experimental validation of the sensor-based calibration method: a) components of the system, and b) setup used for back-to-back comparison of 3D-DIC measurements.

Before performing the 3D-DIC measurement, the two cameras were calibrated using two different procedures. The first one was a traditional image-based calibration approach where several images of a calibration object (i.e., checkerboard) were collected with both cameras simultaneously to extract the intrinsic and extrinsic parameters of the stereovision system. The second approach is based on the developed multi-sensor

system to measure the cameras' relative orientation and position and extract the extrinsic parameters and collect two individual sets of pictures of a calibration object to determine the intrinsic parameters of the system. It should be noted that determining the intrinsic parameters for the sensor-based calibration is a one-time operation and it can be performed at any time as long the focal length, focus, and diaphragm aperture of the two cameras does not change during a test. The extrinsic parameters were computed by recording twenty seconds of data from the IMUs and laser sensor, while the DH transformations [38] were used to convert the computed distance from the laser frame of reference into camera 1 frame of reference. The intrinsic (i.e., optical centers c_x and c_y and focal length f_x and f_y) and extrinsic parameters (i.e., angles α_i, β_i, γ_i, and the components of the translation vector **T**) obtained with the traditional image-based calibration and the proposed sensor-based approach are summarized in **Table II**.

TABLE II
CAMERA PARAMETERS COMPUTED WITH TRADITIONAL AND SENSOR-BASED CALIBRATIONS

		Traditional		Sensor-based	
		Cam 1	**Cam 2**	**Cam 1**	**Cam 2**
Intrinsic Parameters	c_x (pixel)	788.13	825.29	785.85	827.71
	c_y (pixel)	587.68	562.48	588.93	560.82
	f_x (pixel)	2812.70	2824.30	2809.98	2822.10
	f_y (pixel)	2803.20	2811.80	2800.20	2809.70
Extrinsic Parameters (Relative to Camera 1)	α (°)	-0.2708		-0.0290	
	β (°)	0.8319		0.7530	
	γ (°)	19.9886		20.4916	
	T_X (mm)	-627.09		-628.17	
	T_Y (mm)	-3.8106		-6.3882	
	T_Z (mm)	107.47		109.68	

It can be observed that the intrinsic parameters differ on average by ~2 pixels, while the difference in the extrinsic parameters is on the order of millimeters for **T** and fractions of a degree for the Euler angles when the two methods are considered. Once the intrinsic and extrinsic parameters are determined, two calibration files are generated and used to perform 3D-DIC with the open-source software Digital Image Correlation Engine (DICe) [46]. To begin, five images were collected without moving the planar object to determine the noise floor of the measurement (see **Figure 11a** and **Figure 11b**).

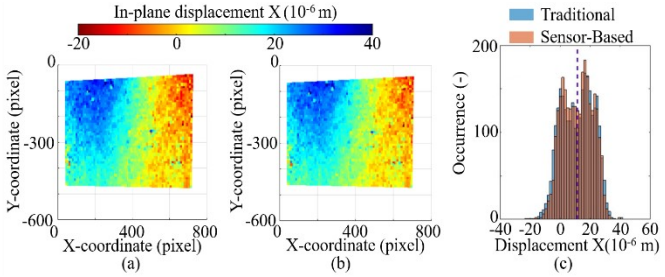


Fig. 11: 3D-DIC noise floor when stereo images were processed using the calibrations obtained with the two methods: a) traditional calibration, b) sensor-based calibration, and c) histograms showing the noise floor distribution with the two calibrations, where the ideal value is zero.

In **Figure 11c**, the displacement histograms showed a high amount of overlap: the K-L Divergence [47] of the two histograms is 3.25, demonstrating a good agreement in the noise floor of the 3D-DIC measurement when the two calibration methods were used. In particular, the in-plane noise floor of the measurement ranges between -20x10^{-6} m and 40x10^{-6} m for both measurements. In addition, the mean value is equal to 11.6x10^{-6} m when the sensor-based calibration is considered compared to a value of 11.4x10^{-6} m when the traditional calibration is used. To finish, the planar object is displaced horizontally towards the left and the displacement results are shown in **Figure 12**.

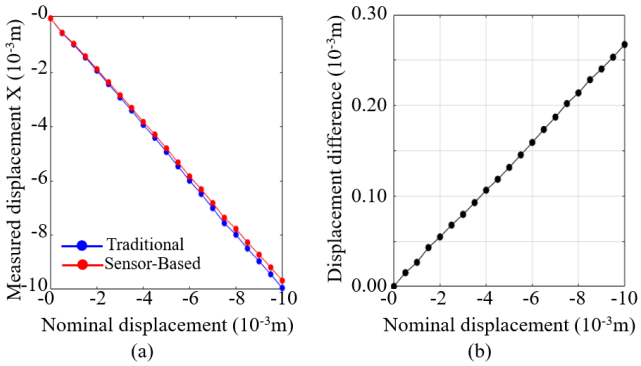


Fig. 12: Back-to-back comparison of 3D-DIC measurements obtained from the two calibrations: a) average measured displacements' values computed at each stage and b) average displacements' absolute difference.

Figure 12a shows the average displacement values computed by each method compared to the nominal displacement applied to the planar object. As observed, both methods yield a linear trend with the sensor-based approach showing an increasing difference between the measured values and the nominal ones, as shown in **Figure 12b**. The relative difference in the displacement computed when the 3D-DIC analysis is performed using the sensor-based calibration and the one computed using the traditional method is lower than 3% (i.e., 260 µm over a displacement of 10 mm) for all displacements, confirming the quality of the proposed multi-sensor system.

VI. CONCLUSIONS

This paper describes the characterization of a multi-sensor system for extracting the extrinsic parameters of a pair of stereo-cameras necessary for performing accurate stereophotogrammetry calibration as an alternative to the traditional image-based procedure. The proposed system is based on one laser sensor to measure the distance between the cameras, and two IMUs to measure the roll, pitch, and yaw angles of the two cameras. A third IMU is also used to overcome the spatial limitations of previously proposed sensor-based calibration methods. Laboratory tests performed to characterize the noise floor of the selected sensors have shown that the precision of the IMUs is on the order of 10^{-3} degree, while the precision of the laser is ~5x10^{-4} m. The IMU and laser's noise floors were then used in an analytical study to understand the effect of sensor noise on the performance degradation of the sensor-based calibration procedure. A simulation of a 3D-DIC measurement on a 0.6x0.8 m planar object showed how the sensor-based method effectively calibrates stereovision cameras yielding reprojection and reconstruction errors below the threshold considered acceptable

for stereophotogrammetry. An experimental evaluation was also performed by comparing the results of a 3D-DIC measurement obtained when the stereo cameras are calibrated using the traditional image-based and the proposed sensor-based calibration. The results of the back-to-back comparison show that the multi-sensor system developed in this study can accurately compute the extrinsic parameters and compute the rigid displacements of a speckled object with an error below 3% when compared to measurements obtained from a traditional calibration. Because the multi-sensor system allows recording the cameras' relative position every time an image is acquired, each image has its own calibration file, making the whole process insensitive to the cameras' relative movement and independent of the size of the targeted structure. As a result, the multi-sensor board described in this paper can extend the applicability of 3D-DIC and open the door for an expedited approach for long-term monitoring of large-scale structures.

REFERENCES

[1] C.-Z. Dong and F. N. Catbas, "A review of computer vision–based structural health monitoring at local and global levels," *Structural Health Monitoring,* vol. 20, no. 2, pp. 692-743, 2021, doi: 10.1177/1475921720935585.

[2] N. A. Valente, A. Sarrafi, Z. Mao, and C. Niezrecki, "Streamlined particle filtering of phase-based magnified videos for quantified operational deflection shapes," *Mechanical Systems and Signal Processing,* vol. 177, p. 109233, 2022/09/01/ 2022, doi: https://doi.org/10.1016/j.ymssp.2022.109233.

[3] D. Feng and M. Q. Feng, "Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection – A review," *Engineering Structures,* vol. 156, pp. 105-117, 2018/02/01/ 2018, doi: https://doi.org/10.1016/j.engstruct.2017.11.018.

[4] C. Niezrecki, J. Baqersad, and A. Sabato, "Digital Image Correlation Techniques for NDE and SHM," in *Handbook of Advanced Nondestructive Evaluation*, N. Ida and N. Meyendorf Eds. Cham: Springer International Publishing, 2019, pp. 1545-1590.

[5] P. F. Luo, Y. J. Chao, M. A. Sutton, and W. H. Peters, "Accurate measurement of three-dimensional deformations in deformable and rigid bodies using computer vision," *Experimental Mechanics,* vol. 33, no. 2, pp. 123-132, 1993/06/01 1993, doi: 10.1007/BF02322488.

[6] T. Schmidt, J. Tyson, and K. Galanulis, "Full-field dynamic displacement and strain measurement using advanced 3D image correlation photogrammetry: Part I," *Experimental Techniques,* vol. 27, pp. 47 - 50, 2006, doi: 10.1111/j.1747-1567.2003.tb00115.x.

[7] D. Reagan, A. Sabato, and C. Niezrecki, "Feasibility of using digital image correlation for unmanned aerial vehicle structural health monitoring of bridges," *Structural Health Monitoring,* vol. 17, no. 5, pp. 1056-1072, 2018, doi: 10.1177/1475921717735326.

[8] M. Shafiei Dizaji, M. Alipour, and D. K. Harris, "Leveraging Full-Field Measurement from 3D Digital Image Correlation for Structural Identification," *Experimental Mechanics,* vol. 58, pp. 1049-1066, 2018.

[9] L. Ngeljaratan and M. A. Moustafa, "Structural health monitoring and seismic response assessment of bridge structures using target-tracking digital image correlation," *Engineering Structures,* vol. 213, p. 110551, 2020/06/15/ 2020, doi: https://doi.org/10.1016/j.engstruct.2020.110551.

[10] A. Sabato and C. Niezrecki, "Feasibility of Digital Image Correlation for railroad tie inspection and ballast support assessment," *Measurement,* vol. 103, pp. 93-105, 02/20 2017, doi: 10.1016/j.measurement.2017.02.024.

[11] M. Kalaitzakis, N. Vitzilaios, D. C. Rizos, and M. A. Sutton, "Drone-Based StereoDIC: System Development, Experimental Validation and Infrastructure Application," *Experimental Mechanics,* vol. 61, no. 6, pp. 981-996, 2021, doi: 10.1007/s11340-021-00710-z.

[12] D. Gorjup, J. Slavič, and M. Boltežar, "Frequency domain triangulation for full-field 3D operating-deflection-shape identification," *Mechanical Systems and Signal Processing,* vol. 133, 11/01 2019, doi: 10.1016/j.ymssp.2019.106287.

[13] A. Khadka, B. Fick, A. Afshar, M. Tavakoli, and J. Baqersad, "Non-contact vibration monitoring of rotating wind turbines using a semi-autonomous UAV," *Mechanical Systems and Signal Processing,* vol. 138, p. 106446, 2020/04/01/ 2020, doi: https://doi.org/10.1016/j.ymssp.2019.106446.

[14] D. Z. Z. Zhang, J. Zhang, and Z Peng, "Improved robust and accurate camera calibration method used for machine vision application," *Optical Engineering,* vol. 47, no. 11, p. 117201, 2008.

[15] F. Remondino and C. Fraser, "Digital camera calibration methods: Considerations and comparisons," *Ine. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.,* vol. 36, 11/30 2005.

[16] Z. Zhang, "Camera Parameters (Intrinsic, Extrinsic)," in *Computer Vision: A Reference Guide*, K. Ikeuchi Ed. Boston, MA: Springer US, 2014, pp. 81-85.

[17] P. Reu, "Hidden Components of DIC: Calibration and Shape Function – Part 1," *Experimental Techniques,* vol. 36, no. 2, pp. 3-5, 2012, doi: https://doi.org/10.1111/j.1747-1567.2012.00821.x.

[18] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 22, no. 11, pp. 1330-1334, 2000, doi: 10.1109/34.888718.

[19] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, "Bundle Adjustment — A Modern Synthesis," Berlin, Heidelberg, 2000: Springer Berlin Heidelberg, in Vision Algorithms: Theory and Practice, pp. 298-372.

[20] J. Zhang, J. Zhu, H. Deng, Z. Chai, M. Ma, and X. Zhong, "Multi-camera calibration method based on a multi-plane stereo target," *Appl. Opt.,* vol. 58, no. 34, pp. 9353-9359, 2019/12/01 2019, doi: 10.1364/AO.58.009353.

[21] D. Solav, K. M. Moerman, A. M. Jaeger, K. Genovese, and H. M. Herr, "MultiDIC: An open-source toolbox for multi-view 3D digital image correlation," *IEEE Access,* vol. 6, pp. 30520-30535, 2018.

[22] P. Sun, N.-G. Lu, M.-L. Dong, B.-X. Yan, and J. Wang, "Simultaneous All-Parameters Calibration and Assessment of a Stereo Camera Pair Using a Scale Bar," *Sensors,* vol. 18, no. 11, 2018, doi: 10.3390/s18113964.

[23] B. Chen, K. Genovese, and B. Pan, "Calibrating large-FOV stereo digital image correlation system using phase targets and epipolar geometry," *Optics and Lasers in Engineering,* vol. 150, p. 106854, 2022/03/01/ 2022, doi: https://doi.org/10.1016/j.optlaseng.2021.106854.

[24] Y. Wang, X. Wang, Z. Wan, and J. Zhang, "A Method for Extrinsic Parameter Calibration of Rotating Binocular Stereo Vision Using a Single Feature Point," *Sensors,* vol. 18, no. 11, 2018, doi: 10.3390/s18113666.

[25] K. Yan, H. Tian, E. Liu, R. Zhao, Y. Hong, and D. Zuo, "A Decoupled Calibration Method for Camera Intrinsic Parameters and Distortion Coefficients," *Mathematical Problems in Engineering,* vol. 2016, p. 1392832, 2016/12/29 2016, doi: 10.1155/2016/1392832.

[26] P. Poozesh, A. Sabato, A. Sarrafi, C. Niezrecki, P. Avitabile, and R. Yarala, "Multicamera measurement system to evaluate the dynamic response of utility-scale wind turbine blades," *Wind Energy,* vol. 23, no. 7, pp. 1619-1639, 2020.

[27] P. Sturm, "Pinhole Camera Model," in *Computer Vision: A Reference Guide*, K. Ikeuchi Ed. Boston, MA: Springer US, 2014, pp. 610-613.

[28] R. Harley and A. Zisserman, "Epipolar geometry and the fundamental matrix,: in Multiple View Geometry in Computer Vision, R. Hartley and A. Zissermann, Eds. 2nd ed. Cambridge, UK: Cambridge Univ. Press, Mar. 2004, pp. 239-259.

[29] M. A. Sutton, J.-J. Orteu, and H. W. Schreier, "Error Estimation in Stereo-Vision," in *Image Correlation for Shape, Motion and Deformation Measurements: Basic Concepts,Theory and Applications*, H. Schreier, J.-J. Orteu, and M. A. Sutton Eds. Boston, MA: Springer US, 2009, pp. 225-226.

[30] W. Feng *et al.*, "Inertial measurement unit aided extrinsic parameters calibration for stereo vision systems," *Optics and Lasers in Engineering,* vol. 134, p. 106252, 2020/11/01/ 2020, doi: 10.1016/j.optlaseng.2020.106252.

[31] A. Sabato, N. A. Valente, and C. Niezrecki, "Development of a Camera Localization System for Three-Dimensional Digital Image Correlation Camera Triangulation," *IEEE Sensors Journal,* vol. 20, no. 19, pp. 11518-11526, 2020, doi: 10.1109/jsen.2020.2997774.

[32] A. Sabato and C. Niezrecki, "Development of an IMU-radar sensor board for three-dimensional digital image correlation camera triangulation," in *Proc.SPIE* 2019, vol. 10972, doi:

10.1117/12.2515081.

[33] D. Kumar, C.-H. Chiang, and Y.-C. Lin, "Experimental vibration analysis of large structures using 3D DIC technique with a novel calibration method," *Journal of Civil Structural Health Monitoring,* 2022/01/25 2022, doi: 10.1007/s13349-022-00549-5.

[34] S. Madgwick, "An efficient orientation filter for inertial and inertial/magnetic sensor arrays," *Report x-io and University of Bristol (UK),* vol. 25, pp. 113-118, 2010.

[35] "LPMS-IG1 Series: High Precision 9-Axis Inertial Measurement Unit (IMU) / AHRS with USB / CAN / RS232 / RS485 Connectivity and Optional GPS Receiver." [Online]. Available: https://lp-research.com/9-axis-imu-with-gps-receiver-series/. [Accessed: 22- Feb- 2022].

[36] "40m Professional Width Measuring Sensor." [Online]. Available: https://www.jrt-measure.com/long-range-distance-sensor/57660459.html. [Accessed: 22- Feb- 2022].

[37] "puA1600-60uc - Basler pulse ", [Online]. Available: https://www.baslerweb.com/en/products/cameras/area-scan-cameras/pulse/pua1600-60uc/. [Accessed: 22- Feb- 2022].

[38] P. I. Corke, "A Simple and Systematic Approach to Assigning Denavit–Hartenberg Parameters," *IEEE Transactions on Robotics,* vol. 23, no. 3, pp. 590-594, 2007, doi: 10.1109/TRO.2007.896765.

[39] F. Bottalico *et al.*, "A sensor-based calibration system for three-dimensional digital image correlation," in *Proc.SPIE*, 2022, vol. 12048, doi: 10.1117/12.2612106.

[40] A. Belhaoua, S. Kohler, and E. Hirsch, "Estimation of 3D reconstruction errors in a stereo-vision system," *Proc SPIE,* vol. 7390, 06/01 2009, doi: 10.1117/12.827346.

[41] Z. Hu, H. Xie, J. Lu, H. Wang, and J. Zhu, "Error evaluation technique for three-dimensional digital image correlation," *Appl. Opt.,* vol. 50, no. 33, pp. 6239-47, Nov 20 2011, doi: 10.1364/AO.50.006239.

[42] K. Yan, R. Zhao, E. Liu, and Y. Ma, "A Robust Fundamental Matrix Estimation Method Based on Epipolar Geometric Error Criterion," *IEEE Access,* vol. 7, pp. 147523-147533, 2019, doi: 10.1109/ACCESS.2019.2946387.

[43] "Evaluating the Accuracy of Single Camera Calibration." [Online]. Available: https://www.mathworks.com/help/vision/ug/evaluating-the-accuracy-of-single-camera-calibration.html. [Accessed: 22- Feb- 2022].

[44] "Rough approximation for measuring accuracy in ARAMIS" [Online]. Available: https://faq.trilion.com/docs/aramis-measuring-accuracy/. [Accessed: 22- Feb- 2022].

[45] A.-S. Poulin-Girard, S. Thibault, and D. Laurendeau, "Influence of camera calibration conditions on the accuracy of 3D reconstruction," *Optics Express,* vol. 24, p. 2678, 02/08 2016, doi: 10.1364/OE.24.002678.

[46] D.Z. Turner, "Digital Image Correlation Engine (DICe) Reference Manual", *Sandia Report*, SAND2015-10606 O, 2015. [Online]. Available: https://dicengine.github.io/dice/ [Accessed: 17-Aug-2022].

[47] J. M. Joyce, "Kullback-Leibler Divergence," in *International Encyclopedia of Statistical Science*, M. Lovric Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 720-722.