Metrol. Meas. Syst., Vol. 29 (2022) No. 1, pp. 75–92
DOI: 10.24425/mms.2022.140028

METROLOGY AND MEASUREMENT SYSTEMS
Index 330930, ISSN 0860-8229
www.metrology.wat.edu.pl

UNCERTAINTY AND ACCURACY OF VISION-BASED TRACKING CONCERNING STEREOPHOTOGRAMMETRY AND NOISE-FLOOR TESTS

Luna Ngeljaratan[1,2)], Mohamed A. Moustafa[1)]

1) *University of Nevada, Department of Civil & Environmental Engineering, Reno, NV 89557, USA,* (*lngeljaratan@nevada.unr.edu*, ✉ *mmoustafa@unr.edu* +1 775 682 7919)
2) *Research Center for Biomaterials, National Research and Innovation Agency, Cibinong Science Center, Jl. Raya Bogor km 46, Cibinong, 16911, Indonesia*

Abstract

This work proposes a systematic assessment of stereophotogrammetry and noise-floor tests to characterize and quantify the uncertainty and accuracy of a vision-based tracking system. Two stereophotogrammetry sets with different configurations, i.e., some images are designed and their sensitivity is quantified based on several assessments. The first assessment evaluates the image coordinates, stereo angle and reconstruction errors resulting from the stereophotogrammetry procedure, and the second assessment expresses the uncertainty from the variance and bias errors measured from the noise-floor test. These two assessments quantify the uncertainty, while the accuracy of the vision-based tracking system is assessed from three quasi-static tests on a small-scaled specimen. The difference in each stereophotogrammetry set and configuration, as indicated by the stereophotogrammetry and noise-floor assessment, leads to a significant result hat the first stereophotogrammetry set measures the RMSE of 3.6 mm while the second set identifies only 1.6 mm of RMSE. The results of this work recommend a careful and systematic assessment of stereophotogrammetry and noise-floor test results to quantify the uncertainty before the real test to achieve a high displacement accuracy of the vision-based tracking system.

Keywords: Stereophotogrammetry, noise-floor, vision-based tracking, displacement, uncertainty, accuracy.

1. Introduction

A vision-based sensor system using the tracking technique, or a *Vision-Tracking* (VT) system, is a non-contact monitoring technique that is widely used for shape and deformation measurements in the field of experimental testing. After 30 years of development, significant improvements have been attained not only in measurement accuracy and computational efficiency but also in laboratory and field applications [1–7]. However, even though major efforts have been dedicated to applying VT in *structural health monitoring* (SHM), at least from the academic point of view, the comprehensive standard of practice or guidelines for monitoring civil infrastructure systems

Article history: received August 13, 2021; revised November 17, 2021; accepted December 3, 2021; available online December 30, 2021.

is not yet available. The method to assess the uncertainty and accuracy level of VT systems is not yet standardized and only a few guidelines are available for the specific implementation of computer vision algorithms in laboratory tests [8]. Also, even though considerable theoretical work has been done in both image correlation and stereovision, only very few studies consider large-scale experimental validation of the VT system [6, 9–12].

Stereophotogrammetry is a technique that measures a geometric dimension of an object in spatial position by reconstructing 3D targets or points found on a set of images. Reconstruction of 3D points is a crucial task in stereophotogrammetry and requires the intrinsic and extrinsic parameters of each camera to be precisely and accurately calibrated. Therefore, unlike classical measurement techniques, the uncertainty and accuracy of a VT system depend strongly on the quality of stereophotogrammetry and camera calibration. Various parameters are closely related to that quality such as the pattern form, air refraction, distortion of the image, and the image processing algorithm [13]. The patterns can be artificial or natural points or targets on the object in which an improper form may increase the error of VT system measurement. When the monitoring is conducted under different air temperature conditions, the light propagation path will change causing the refraction deviation to increase significantly. Image distortion may be caused by the optical or electronic system or image digitizing while the accuracy of image processing directly depends on the resolution. The sampling frequency, types of lightings or camera movements are also factors that affect the accuracy of a VT system [1, 6, 10, 14]. Since these factors are already considered in the previous studies, this work focuses more on a direct quantification of uncertainty and accuracy of the stereophotogrammetry and camera calibration.

Prior to this work, several studies already attempted to investigate the stereophotogrammetry and calibration uncertainties with several recommendations to improve their accuracy. A study by Zhu *et al.* [15] specified the extracted features from images as the source of calibration uncertainty with a model proposed to design a multiple-stage calibration algorithm. The camera intrinsic parameters were also modelled with a numerical solution provided by the Monte Carlo approach [16–18]. The uncertainty of extrinsic parameters in terms of the detected feature position, orientation, and angles of the detected features on the measured object was also estimated and successfully improved the accuracy with margin error within 1.3–1.8 mm [19] and uncertainty of 27–49 mm using the global method [20]. Since the global calibration method alone only provides a conservative approach in measurement [21]; a local method [20] was also proposed and improved the accuracy to 2 mm. A robust approach using local multi-cameras and non-linear optimization has also been proposed [22, 23] but the number of cameras may increase the false matches and decrease the S/N ratio [24]. Other methods such as modelling the motion blur [25], the back projection process [26], trained neural network [27], rotating axis calibration [28, 29] or pointing error methods [30] can also model the uncertainty and improve the vision-based system accuracy.

Only a few studies conducted so far focused on the verification and validation of VT system data concerning other instrumentation data from a large-scale structure test. A two-span bridge model was tested under amplified seismic loads and measured RMSE of 0.317 in between the VT system and a string potentiometer [6]. Concrete and steel bridges were also tested under seismic excitation in a laboratory test and the system identification measured their modal characteristics with a variation of frequency between 0.01–0.03 Hz and a damping ratio of 0.07–0.5% [12]. A VT system using a single camera model was also tested on a railway bridge with an RMS of 0.2 mm based on a camera-to-target distance of 6.9 m [31]. Other studies also investigated the accuracy of the VT system in monitoring civil structures in the field. However, the results were not validated using other instrumentations [4] or the accuracy was not reported [5].

As compared to prior works, this study emphasizes quantifying the vision-based tracking uncertainty and accuracy by exploring the sensitivity of the existing stereophotogrammetry technique and exploring a noise-floor test as an alternative tool for preliminary checks. Measuring the static and dynamic behavior of large-scale civil structures in high accuracy through vision-based tracking SHM is the final goal. Therefore, recommending an experimental approach to minimize the uncertainty and improve the accuracy is considered rather than proposing another theoretical work or hybrid method that has already been complimented previously in many studies prior to this work. The goal of this study is to characterize and quantify the uncertainty and accuracy of vision-based tracking systems through a systematic assessment of stereophotogrammetry and noise-floor tests. The paper first defines VT system stereophotogrammetry with the associated error source. Next, the validation tests and results are discussed in full detail. Three quasi-static tests provide data for VT validation and the uncertainty margin as well as the true displacement error are estimated based on the proposed assessment.

2. Method

2.1. Stereophotogrammetry procedure and uncertainty

An overview of a VT system application concerning camera self-calibration and stereophotogrammetry is shown in Fig. 1. The object is first prepared by determining the monitored area of interest, monitoring station, and setting the camera. The features to track can be natural or artificial, in which a circular template is preferable as it is faster to identify using an ellipse finding algorithm. For 3D measurements, two cameras or more are calibrated, starting from collecting a set of object images from different positions, orientations, and angles. The intrinsic and extrinsic parameters of the cameras are determined from the calibration with options ranging from non-linear to multiple-stage calibration techniques. The coordinate transformation from image coordinates (2D, pixels) to an object coordinate system (3D, mm) is estimated through the stereophotogrammetry process that relies on the triangulation of a circular template found in those collected images. Therefore, those features, or templates, should be uniformly distributed on the object and adequately captured in each image.

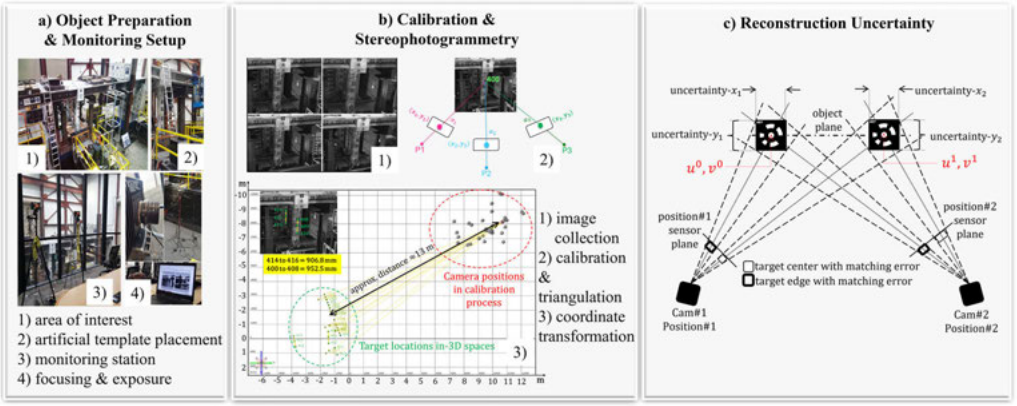

Fig. 1. a, b) General procedure of VT system application for laboratory and field testing using camera self-calibration and stereophotogrammetry techniques; c) Reconstruction uncertainty in stereophotogrammetry.

Stereophotogrammetry in this study uses physical targets with an example shown in Fig. 1c. Several options of physical targets are available. However, this study implements the template design of Schneider [33] who uses a white circle surrounded by white rings printed on a black background. The selection is based on the ease of identification of each target as the configuration of the white ring is unique and associated with a specific numbering. The stereophotogrammetry applies the camera self-calibration technique [32] using two cameras, so that several images are captured by moving the cameras from different locations. In consequence, the targets in each image are captured in different spatial orientations. In large area stereophotogrammetry conducted in this study, the center of the targets as shown in Fig. 1c is initially found using the ellipse finding algorithm [33]. Then, the spatial orientation of the 2D image is computed so that all targets must be seen in at least three images to compute the point rays. The coordinates of the target in space are identified using triangulation and bundle adjustment. Bundle here means that the point rays of each target should intersect to complete the triangulation process. The movement of each target during the test is measured using the pattern matching method that holds the principle of finding matching interesting regions with precise locations between the pixel subsets in both images on the cameras using a correlation function. Consequently, deformation measurements are generated between different images using the convolution relationship with the transformation matrix. The details of the stereophotogrammetry conducted in this study as well as the camera calibration procedures were already presented and can be found in our previous works [6, 10, 11].

The camera model in this study is the pinhole camera one [34, 36] so that the transformation of the target world coordinates XYZ from the pixel coordinates uv is established through the relationship between camera intrinsic matrix or internal parameters A and extrinsic matrix or external parameters as shown in (1). The focal length f_x, f_y and camera center c_x, c_y form the intrinsic matrix while the extrinsic matrix consists of rotation matrix R and translation matrix T.

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A\,[R\;T] = \begin{bmatrix} f_x & f_s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}. \tag{1}$$

As shown in Fig. 1c, each camera has its own pixel coordinate system $\left(u^i, v^i\right)$ with $i = 0.1$ corresponding to the two cameras with superscripts 0 and 1 used to denote camera #1 and #2, respectively. Since each camera is calibrated separately, the projection of a spatial point onto the camera sensor planes in one image can be written as per (2) [34].

$$\begin{cases} u^0 = f_x^0 \dfrac{r_{11}^0 X_w + r_{12}^0 Y_w + r_{13}^0 Z_w + t_x^0}{r_{31}^0 X_w + r_{32}^0 Y_w + r_{33}^0 Z_w + t_z^0} + f_s^0 \dfrac{r_{21}^0 X_w + r_{22}^0 Y_w + r_{23}^0 Z_w + t_y^0}{r_{31}^0 X_w + r_{32}^0 Y_w + r_{33}^0 Z_w + t_z^0} + c_x^0 \\ v^0 = f_y^0 \dfrac{r_{21}^0 X_w + r_{22}^0 Y_w + r_{23}^0 Z_w + t_y^0}{r_{31}^0 X_w + r_{32}^0 Y_w + r_{33}^0 Z_w + t_z^0} + c_y^0 \\ u^1 = f_x^1 \dfrac{r_{11}^1 X_w + r_{12}^1 Y_w + r_{13}^1 Z_w + t_x^1}{r_{31}^0 X_w + r_{32}^0 Y_w + r_{33}^0 Z_w + t_z^0} + f_s^1 \dfrac{r_{21}^1 X_w + r_{22}^1 Y_w + r_{23}^1 Z_w + t_y^1}{r_{31}^1 X_w + r_{32}^1 Y_w + r_{33}^1 Z_w + t_z^1} + c_x^1 \\ v^1 = f_y^1 \dfrac{r_{21}^1 X_w + r_{22}^1 Y_w + r_{23}^1 Z_w + t_y^1}{r_{31}^1 X_w + r_{32}^1 Y_w + r_{33}^1 Z_w + t_z^1} + c_y^1 \end{cases}. \tag{2}$$

If the reference sensor plane location is assumed to be error-free, then the center-point pixel location of camera #1 $\left(u^0, v^0\right)$ is exact and camera #2 center-point pixel location $\left(u^1, v^1\right)$ shown

in Fig. 1b implies the optimally matched subset. However, for a large-area stereophotogrammetry, this assumption becomes invalid. A larger target dimension is required so more complex stereophotogrammetry is necessary [3]. Reu [21] indicated that if the stereophotogrammetry were successful, then the matching error was the major source of the 3D position error for a larger target as illustrated in Fig. 1b. Since the pixel coordinates of both cameras do not strictly follow Gaussian distribution with a zero-mean value, both uncertainties in horizontal and vertical positions, x and y, are present in large-area stereophotogrammetry. This uncertainty results in a reconstruction error that is strongly related to the stereo angle, focal length, working distance, and other camera parameters. The coordinate matching error, if exists, can also be the main source of 2D and 3D position errors [21]. As for the uncertainty in the direction towards the camera, *i.e.*, out-of-plane or Z direction, a comprehensive study [35] revealed that the out-of-plane movement was unavoidable during the loading process in the in-plane directions and became the major source of uncertainty in VT systems. The distance from the camera to the object should be increased for more accurate results when the measurement is conducted using a single camera.; However, optimizing the camera calibration was the only offered solution to minimize the error in the Z direction for multiple camera systems.

2.2. Assessment and sensitivity analysis

The required images in one stereophotogrammetry set can vary based on the measured FOV, but it is recommended in [8] to consider at least 8 and up to 100 images in one set to achieve higher measurement accuracy. Meanwhile, there are no specific requirements in common practice guides [8] either for the number of stereophotogrammetry sets or how many times the VT system should be calibrated to minimize the measurement error. The measurement accuracy can only be tested by running a laboratory or field experiment and comparing the VT system results with other types of reference sensor measurement. If the error of the VT system with respect to the reference sensors is high, then the stereophotogrammetry result should be rejected and the cameras must be recalibrated, only if the VT system is still preferred in the monitoring. There is no instant approach to assess the accuracy of the VT system before conducting the real tests or without comparing VT results with other instrumentations. Meanwhile, running more than one round of stereophotogrammetry is a quick and easy option. Therefore, this study investigates a practical solution *i.e.* collecting two stereophotogrammetry sets with each set consisting of different calibration images. Running two rounds of stereophotogrammetry delivers two preliminary results then are verified in a noise-floor test before conducting the real tests.

In this study two sensitivity analysis parameters, *i.e.* the number of stereophotogrammetry sets and the sensitivity of the image quantities are designed. Two stereophotogrammetry sets, *i.e.* Set#1 and Set#2 with five configurations of different image quantities, *i.e.* A-E, for each set are shown in Fig. 2a. The configuration is defined as the total number of calibration images processed in each stereophotogrammetry set as follows: 10 images for configuration A, 18 images for configuration B, *etc.* In the rest of the paper, the stereophotogrammetry set is mentioned along with a specific configuration. For example, 1A refers to stereophotogrammetry Set#1 with 10 images corresponding to configuration A, *etc.* The uncertainty for stereophotogrammetry and the noise-floor test only checks the sensitivity of Set#1E and Set#2E using 32 images for simplicity. As for quantifying the accuracy from the laboratory tests, all configurations are considered and the results are reported.

The assessment parameters with their associated method to quantify the uncertainty and accuracy of the vison-based system as proposed in this study are shown in Fig. 2a. The uncertainties of the VT system are evaluated from stereophotogrammetry and noise-floor tests while the accuracy

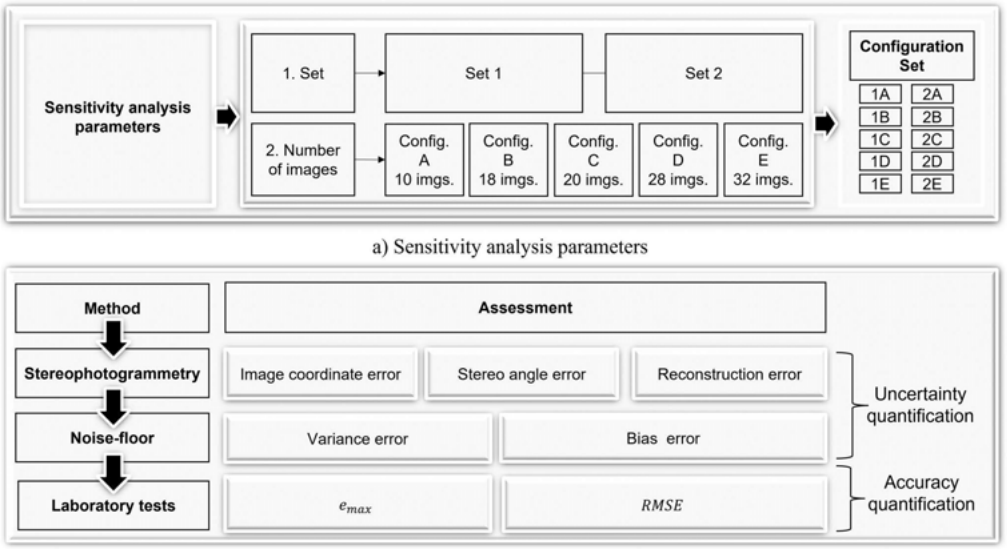

Fig. 2. a) Stereophotogrammetry sets and configurations; b) Practical method and assessment of a vision-based system.

is assessed through conducting a laboratory test using quasi-static loads. In the stereophotogrammetry method, three parameters, *i.e.*, image coordinate, stereo angle, and reconstruction errors are assessed using (3) and (4). Let us assume that target i is captured by camera#1 and camera#2 and the target in each camera is in a horizontal (x_{1i}, x_{2i}) and a vertical (y_{1i}, y_{2i}) position in 2D-image coordinates (pixel). Each target is also oriented toward each camera as $\theta_{1i}\theta_{2i}$ The errors for each target in image coordinates, *i.e.*, horizontal error, e_{x_i} and vertical error, e_{y_i} as well as stereo angle error, e_{θ_i}, are computed using (3):

$$e_{x_i} = x_{1i} - x_{2i}; \qquad e_{y_i} = y_{1i} - y_{2i}; \qquad e_{\theta_i} = \theta_{1i} - \theta_{2i}\,. \tag{3}$$

The reconstruction error is measured from the coordinate transformation result in 3D-object coordinates (mm). To simplify, the reconstruction error is expressed as the distance error between each target on the specimen measured from the stereophotogrammetry in object coordinates, *i.e.* mm, as compared to the real distance that is measured manually. If we assume the vertical position in object coordinate of target iY_i, and target jY_j, with the real distance measurement, $D_{i,j}$, the reconstruction error, E is measured by (4):

$$E = \frac{\left|Y_i - Y_j\right|}{D_{i,j}} \times 100\%. \tag{4}$$

The assessment in the noise-floor test using variance error and bias error is explained in detail in section 3.3. As for the laboratory test, the accuracy quantification is assessed for the displacement value using the maximum error, $e_{\max}$ and *Root Mean Square Error* (RMSE) between the VT system result with respect to the calibrated string-potentiometer result.

2.3. Noise-floor test

In our study the noise-floor test was conducted before the laboratory test by recording images of the unloaded specimen within a specified duration. The variance and bias errors are the parameters used to further investigate the uncertainty of the stereophotogrammetry sets. Variance error refers to a random error centered with a mean around the true value of displacement. The main sources of variance error are camera noise and the matching error as the result of the correlation process. Bias error is often difficult to quantify the noise-floor test because the true value of displacement is known only through running a real test with any applied type of loading. Nonetheless, the bias error of the displacement can be quantified from the mean of the displacement from the noise-floor test as proposed in this study. These errors are evaluated for each loading direction. The out-of-plane direction or the direction towards the camera focus is usually measured to identify the noise-floor since the average pixel deviation is usually larger in this direction (0.03 pixel) as compared to the other directions (0.01 pixel) [36, 37]. However, the evaluation should also be conducted in the direction of the applied load to ensure that the noise sensitivity in the monitored direction is very close to zero.

2.4. Test setup

The proposed method is experimentally evaluated through laboratory validation using three quasi-static tests. The model is an aluminum block specifically constructed for this study with the dimensions 100 mm × 10 mm × 900 mm as shown in Fig. 3. The approximated *field of view* (FOV) for this test was 2.5 m × 1.2 m × 0.05 m. The test setup is shown in Fig. 3 along with VT targets on one side of the model and the shake table. Eleven targets (numbered from 399 to 409) were attached to the specimen and six targets (numbered 410 to 416) were distributed on the shake table. For the best and widest FOV for the test, the monitoring station was located approximately 3.8 m from the specimen.

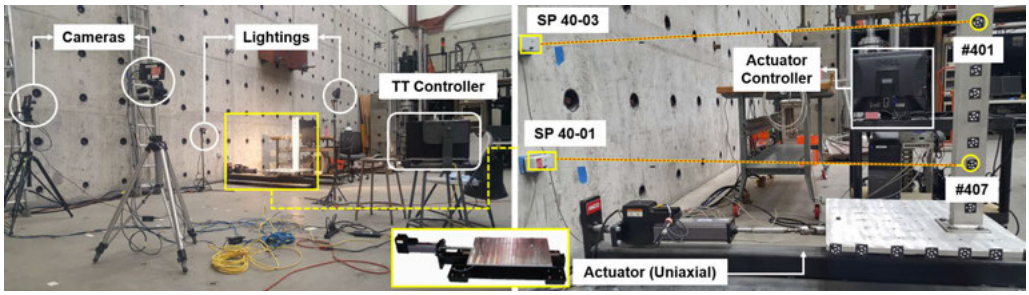

Fig. 3. Monitoring setup (left) and the details of the specimen with the targets used for comparison with the calibrated mechanical sensors (right).

Two high-speed cameras with specifications given in Table 1 are used as the main component of the VT system. Full 5MB pixel resolution of the cameras was used with monochrome types and 8-bit images. The target pattern was a white circle with 3.5 mm diameter and was surrounded by white rings and then printed on black background, like the example in Fig. 1. For both noise-floor tests and validation tests, monitoring was conducted with a sampling rate of only 32 Hz and a global electronic shutter of 3000 ms. Two NIST-calibrated string potentiometers, *i.e.* an SP 40-03 (top channel) and SP 40-01 (bottom channel) were connected to the specimen as shown in Fig. 3. SP 40-03 was connected to the specimen close to target #401 (top) and SP 40-01 was

connected close to target #407 (bottom). The first test was run with a prescribed displacement input to the small shake table actuator, while the second and third tests were conducted manually by setting the table free and pushing the specimen. The direction of the applied load with respect to the camera orientation was in the horizontal in-plane direction. The uncertainty quantification is assessed on the vertical as well as the out-of-plane direction, while the accuracy is only estimated in the applied direction of loading.

Table 1. Specifications of the vision-based tracking system.

Binary	8-bit	**Lens**	35 mm	**Shutter**	Global electronic shutter from 3 μs to 41.654 ms
Communication port	USB 2.0 device port (micro-B), Ethernet (10/100/1000Base-T)	**Resolution (pix.)**	2560 × 2048	**Target size (radius)**	3.5 mm
File format	TIFF	**Sensor**	CMOS	**Type**	Monochrome

3. Results and discussion

3.1. Camera self-calibration

The two stereophotogrammetry sets, *i.e.* Set#1 and Set#2 as shown in Fig. 2 consisted of 32 images (16 images of each camera) that had been captured in different spatial orientation, positions, and angles using the camera self-calibration technique. All targets were well-distributed in each image to enable the least-squares adjustment for estimating the camera internal parameters and to converge the bundle adjustment process. The camera self-calibration results are given in Table 2 for Set#1E and Set#2E associated with a total of 32 processed images. The table shows the locations of the principal point in the x and y directions $(u_0,\ v_0)$, as well as the radial symmetric, asymmetric, and tangential distortions $A_1,\ A_2,\ A_3,\ B_1,\ B_2$ respectively. They are given for Cam#1 and Cam#2 as each camera is calibrated separately.

Table 2. Camera self-calibration results from stereophotogrammetry Set#1E and Set#2E using 32 images.

Set#1E	$u_0,\ v_0$ **(pix)**	A_1	A_2	A_3	B_1	B_2
Cam#1	13.7, –8.2	-3.5×10^{-9}	-1.1×10^{-15}	1.7×10^{-22}	-1.6×10^{-8}	-1.1×10^{-7}
Cam#2	2.8, 13.2	-5.1×10^{-9}	7.0×10^{-17}	-1.7×10^{-22}	-1.6×10^{-7}	-1.3×10^{-7}
Set#2E	$u_0,\ v_0$ **(pix)**	A_1	A_2	A_3	B_1	B_2
Cam#1	13.7, –8.9	-3.6×10^{-9}	-1.0×10^{-15}	1.0×10^{-22}	-1.6×10^{-8}	-1.1×10^{-7}
Cam#2	4.2,14.9	-5.1×10^{-9}	5.8×10^{-18}	-1.3×10^{-22}	-1.6×10^{-7}	-1.3×10^{-7}

3.2. Uncertainty from stereophotogrammetry set

3.2.1. Image coordinate error

The image coordinate error for both stereophotogrammetry sets is first computed using (3) and measured based on the position of targets 399–409. As shown in Fig. 4 the means of the two camera systems lie within their respective error bars. For the horizontal error, e_x a significant

error of target coordinate measured by Set#1E is already observed near the image center, while it is very small for stereophotogrammetry Set#2E. An approximate error of 30 pixels is observed on target 399 (top specimen) that increases linearly up to an error of 60 pixels on target 409 (bottom specimen). While for stereophotogrammetry Set#2E, the horizontal error is very small for the targets near the image center, it increases gradually for those away from the image center. In the vertical direction, both stereophotogrammetry sets yield identical errors e_y and trends. This assessment shows that the image coordinate error is present in both sets with a smaller error for Set#2, and the error in horizontal direction shows more variations than the vertical position between the two sets.

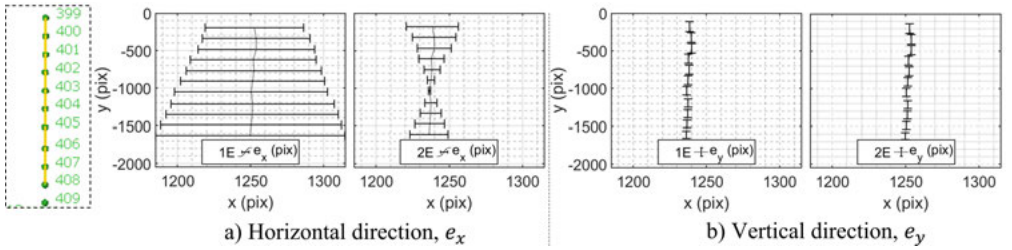

Fig. 4. Image coordinate error of each target in the horizontal, e_x and vertical, e_y directions measured from stereophotogrammetry Set#1E and Set#2E.

3.2.2. Stereo angle error

For the second assessment, the statistical error on the stereo angle for each target is computed using (3) and the results are shown in Table 3 for stereophotogrammetry Set#1E and Set#2E. The stereo angle error is present in both stereophotogrammetry sets, however, the values are less than 1.5° which is considered small. Similar to the horizontal position error observed previously, the stereo angle error also propagates from the top target #399 to the bottom target #409 for both sets. However, the difference of stereo angle error between both sets is insufficient to determine which set provides larger uncertainty. As measured by the average error in Table 3, the stereo angle error is only slightly higher for stereophotogrammetry Set#1E as compared to Set#2E.

Table 3. Stereo angle error, $\mathbf{e}_{\theta_i}$ (°) of each target measured from stereophotogrammetry Set#1E and Set#2E.

Set	**Target, *i***											**Avg.**
	399	400	401	402	403	404	405	406	407	408	409	
1E	0.09	0.36	0.37	0.79	0.70	1.02	1.04	1.06	0.96	1.20	1.09	0.79
2E	0.19	0.32	0.39	0.72	0.63	0.80	0.97	1.09	0.99	1.25	1.16	0.77

3.2.3. Reconstruction error

The last assessment in quantifying the uncertainty of stereophotogrammetry is estimating the reconstruction error of each target in object coordinate. The results are shown in Table 4, which demonstrates the configuration in each set is not sensitive to the reconstruction error. Using more images does not necessarily improve the accuracy of each stereophotogrammetry set. Meanwhile, designing more than only one stereophotogrammetry set shows a possibility to avoid higher reconstruction error as the stereophotogrammetry Set#2 resulted in lower uncertainty and significantly lower error when compared to Set#1 for all configurations.

Table 4. Reconstruction error expressed as distance error obtained from Set#1 and Set#2.

Set	Real distance (mm)									
	399–400	400–401	401–402	402–403	403–404	404–405	405–406	406–407	407–408	408–409
	78	79	81	82	78	80	93	80	79	79
	Error, E (%)									
	399–400	400–401	401–402	402–403	403–404	404–405	405–406	406–407	407–408	408–409
1A	12.5	12.1	17.3	12.6	12.9	16.9	16.3	10.5	15.1	12.3
2A	*2.0*	*2.4*	*4.0*	*0.6*	*0.2*	*0.1*	*1.0*	*0.2*	*0.6*	*0.9*
1B	11.1	10.9	16.7	13.9	12.6	13.2	16.0	13.3	13.4	12.2
2B	*1.8*	*2.3*	*3.9*	*0.6*	*0.2*	*0.1*	*1.0*	*0.2*	*0.5*	*0.8*
1C	12.5	12.1	17.3	12.6	12.9	16.6	17.6	9.5	15.1	12.3
2C	*1.9*	*2.4*	*4.0*	*0.6*	*0.2*	*0.1*	*1.0*	*0.2*	*0.6*	*0.9*
1D	9.7	9.5	14.5	12.1	10.9	11.5	14.0	11.5	11.7	10.6
2D	*1.9*	*2.4*	*4.0*	*0.6*	*0.2*	*0.1*	*1.0*	*0.2*	*0.6*	*0.9*
1E	12.5	12.3	18.2	15.4	14.0	14.6	17.8	14.7	14.9	13.6
2E	*2.0*	*2.5*	*3.9*	*0.5*	*0.3*	*0.0*	*0.9*	*0.1*	*0.5*	*1.0*

The uncertainties from stereophotogrammetry reveal that conducting more than one set may lead to a noticeable difference in image coordinate and reconstruction errors. The result shows that the stereophotogrammetry Set#2 generates a smaller error as compared to Set#1. Meanwhile, the stereo angle error is relatively small and not significant when both sets are compared. The number of images processed in stereophotogrammetry did not contribute to decreasing the error either. Thus, it is desirable to further check whether a consistent difference between the two stereophotogrammetry sets is also observed in the noise-floor test.

3.3. Uncertainty from the noise-floor test

3.3.1. Variance error

The noise-floor test was conducted by recording images of the specimen without applying any load within a 16-sec time window (480 images from each camera). Set#1E and Set#2E were used to process similar noise-floor test images to generate a displacement response history for each set. Variance error is expressed as spatial and temporal variations that are computed from the displacement of noise-floor test images and take both approaches into consideration. First, the standard deviation of all targets in each time step is computed. Then, it is plotted for overall test duration with the results shown in Fig. 5a and 5b. Second, the standard deviation of each target for the entire load step is computed separately with the results shown in Fig. 6. Note that the standard deviation shown in Fig. 6 is computed for each camera in both sets, *e.g.* 2E-1 is the standard deviation of all targets measured from Set#2E camera#1, *etc*. Spatial variations from each set, $\mu_{1E,S}$ and $\mu_{2E,S}$ are simply computed by taking the average of the SD error in Fig. 5 and the values are shown also in Fig. 5. The temporal variations, $\mu_{1E,T}$ and $\mu_{2E,T}$, are also estimated by measuring the average of standard deviation of all targets given in Fig. 6. These two methods for quantifying variance error are applied for all possible loading cases, *i.e.* horizontal (Y) vertical (Z) and out-of-plane (X) directions.

Metrol. Meas. Syst., Vol. 29 (2022), No. 1, pp. 75–92
DOI: 10.24425/mms.2022.140028

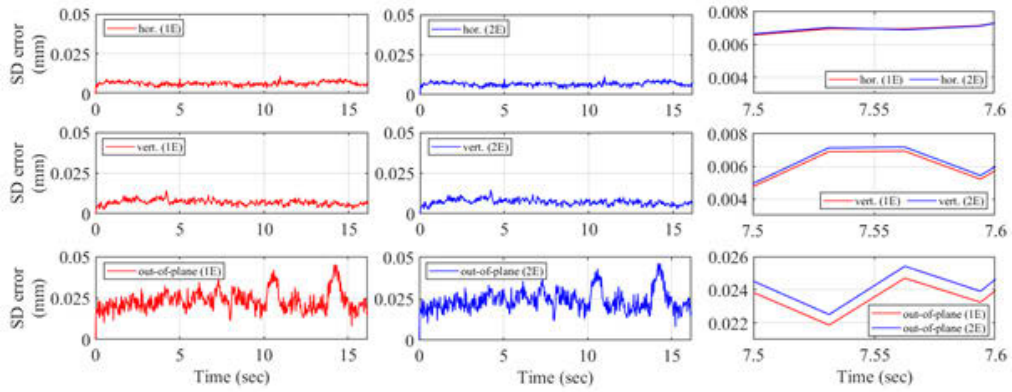

Fig. 5. The standard deviation computed from all targets in each time step and spatial variations of stereophotogrammetry a) Set#1E, $\mu_{1E,S}$, b) Set#2E, $\mu_{2E,S}$ and c) Zoomed-view for clear observation of each set standard deviation error and the difference between the two sets.

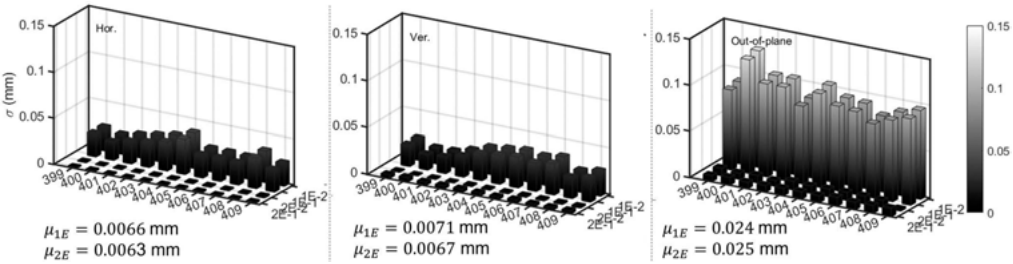

Fig. 6. The standard deviation of all targets and the temporal variations of stereophotogrammetry Set#1E, $\mu_{1E,T}$ and Set#2E, $\mu_{2E,T}$.

The variance error shown in Figs. 5 and 6 indicates that the uncertainty increases in the out-of-plane direction for both stereophotogrammetry sets. These values are the minimum uncertainties of the in-plane and out-of-plane displacements for both sets indicating that the displacement error from the real tests is expected to be larger than these values. The significant observation from Fig. 5 and Fig. 6 is that the uncertainty is larger for Set#1E compared to Set#2E in all possible loading cases. It confirms to the results from the previous section *i.e.* Set#2E eventually produces lower uncertainty than Set#1E.

3.3.2. Bias error

The mean bias error as shown in Fig. 7 is computed from the mean displacement of all targets measured for the entire time step. A similar trend of the variance error is also observed in the bias error *i.e.* the out-of-plane uncertainty is larger as compared to the in-plane direction. Comparing the variance and bias error in all directions, the bias error is approximately two times larger than the variance error. The mean displacement varying over time in Fig. 7 indicates that the bias error exists in both stereophotogrammetry sets. Similarly to the variance error, the bias error observed from the noise-floor test also does not reflect the bias of the true displacement and it is estimated from a real test as described in the next section. They rather estimate the minimum uncertainty from both stereophotogrammetry sets.

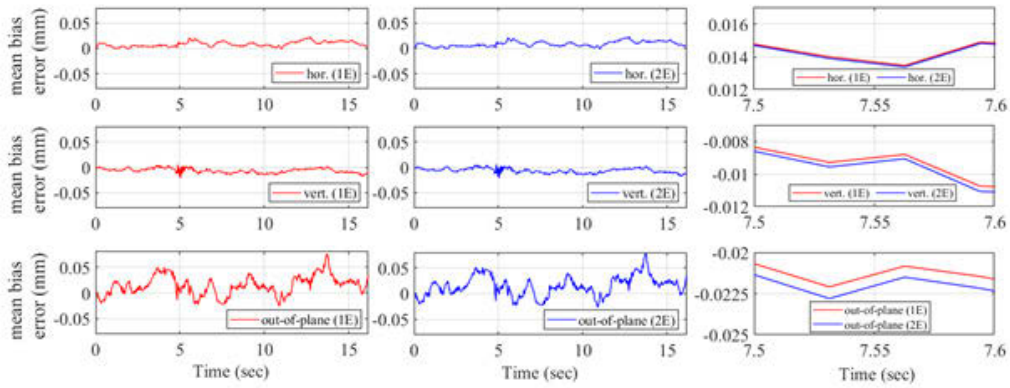

Fig. 7. Bias error of stereophotogrammetry a) Set#1E, b) Set#2E, and c) Zoomed-view for clear observation of each set mean bias error and the difference between the two sets.

The uncertainty quantification from stereophotogrammetry and the noise-floor test shows that they may affect the measurement quality so that the stereophotogrammetry Set#2 shows lower uncertainty when compared to Set#1. A validation test is needed to further confirm this observation and to provide a better understanding of estimating the VT displacement accuracy from different stereophotogrammetry sets.

3.4. Measurement accuracy from the validation test

This section presents the comparison and error analysis of VT system displacement measurements with respect to calibrated string potentiometers from three different quasi-static tests. The example of the displacement history is given in Fig. 8 for all tests. The associated error propagation computed from all configurations is also plotted in Fig. 9. The details of measurement error computed from all tests are given in Table 5.

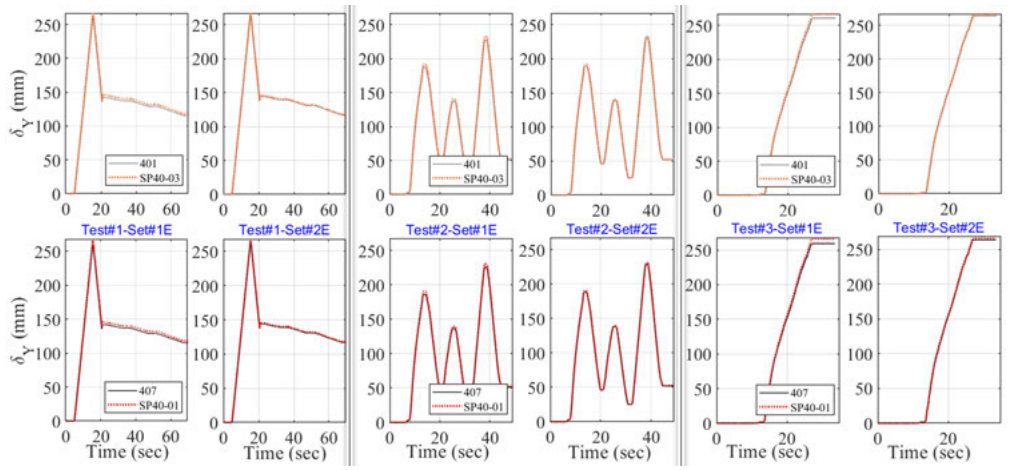

Fig. 8. Comparison of displacement history between string potentiometers and vision-based system targets using stereophotogrammetry Set#1E and Set#2E measured from Test#1, 2, and 3.

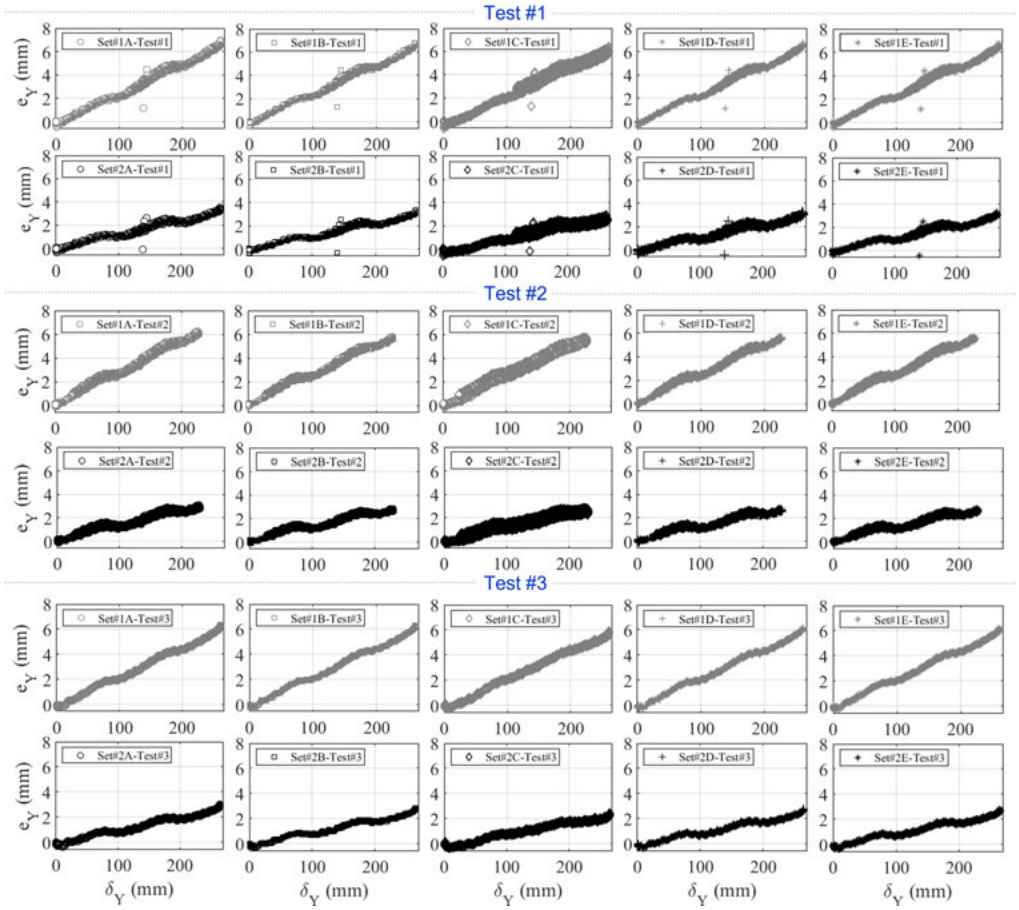

Fig. 9. Relative displacement error, $\mathbf{e_Y}$ with respect to measured displacement of the vision-based system, $\boldsymbol{\delta_Y}$ of Set#1 and Set#2 for configurations A-E.

Figure 8 shows a displacement comparison between target#401 versus SP 40-03 near the top of the specimen, and target#407 versus SP 40-01 near the bottom of the specimen when results from Set#1E and Set#2E are used, respectively. The first observation is that there is no significant difference between the two sets in measuring the displacement as compared to the string potentiometer measurement. More details are shown only when the difference is plotted against the propagation of the displacement amplitude in Fig. 9. A bias error has occurred in both stereophotogrammetry sets and configurations that increased systematically together the applied displacement. Second, it is observed that the error is insensitive to different configurations A-E associated with the number of images used in each set. For example, as shown in Table 5, the maximum error from Test#1 computed using Set#1A and Set#1E shows a very small variation, *i.e.* 2.6% and 2.5%, respectively. The most significant difference is again observed when the results are compared between the two stereophotogrammetry sets. Stereophotogrammetry Set#2 features considerably smaller errors than results based on Set#1. As shown in Table 5 and from all three tests, the maximum error of Set#1 is computed as 2.6% with RMSE of 3.3 mm while for Set#2, a smaller maximum error and RMSE are observed as 1.3% and 1.6 mm, respectively.

Table 5. Measurement error from validation tests of TT displacement with respect to string potentiometers when the DIC was processed using different stereophotogrammetry sets and configurations.

Channels	**Error**		**Test #1**									
			Config. A		**Config. B**		**Config. C**		**Config. D**		**Config. E**	
			Set 1	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**
VT#401 vs. SP 40-03	$e_{\max}$	mm	7.0	3.5	6.8	3.4	6.7	.5	6.7	3.3	6.7	3.3
		%	2.6	1.3	2.6	1.3	2.5	1.3	2.5	1.2	2.5	1.2
	RMSE	mm	3.3	1.6	3.3	1.5	3.2	1.6	3.2	1.5	3.2	1.5
VT# 407 vs. SP 40-01	$e_{\max}$	mm	6.5	3.1	6.5	2.9	6.4	2.9	6.4	2.8	6.4	2.8
		%	2.4	1.2	2.4	1.1	2.4	1.1	2.4	1.1	2.4	1.1
	RMSE	mm	3.3	1.6	3.3	1.5	3.2	1.5	3.2	1.4	3.2	1.4
Channels	**Error**		**Test #2**									
			Config. A		**Config. B**		**Config. C**		**Config. D**		**Config. E**	
			Set 1	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**
VT#401 vs. SP 40-03	$e_{\max}$	mm	6.3	3.1	5.9	2.9	5.7	2.9	5.7	2.8	5.7	2.8
		%	2.6	1.3	2.5	1.3	2.5	1.2	2.5	1.2	2.5	1.2
	RMSE	mm	3.1	1.6	2.9	1.5	2.8	1.5	2.8	1.4	2.8	1.4
VT#407 vs.SP 40-01	$e_{\max}$	mm	6.3	3.1	5.9	2.9	5.7	2.9	5.8	2.8	5.8	2.8
		%	2.7	1.3	2.5	1.3	2.5	1.2	2.5	1.2	2.5	1.2
	RMSE	mm	3.1	1.5	2.9	1.4	2.8	1.4	2.8	1.4	2.8	1.4
Channels	**Error**		**Test #3**									
			Config. A		**Config. B**		**Config. C**		**Config. D**		**Config. E**	
			Set 1	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**
VT#401vs.SP 40-03	$e_{\max}$	mm	6.3	3.0	6.3	2.9	6.2	2.9	6.2	2.8	6.2	2.8
		%	2.4	1.1	2.4	1.1	2.3	1.1	2.3	1.0	2.3	1.0
	RMSE	mm	3.6	1.7	3.6	1.6	3.6	1.6	3.6	1.5	3.6	1.5
VT#407 vs.SP 40-01	$e_{\max}$	mm	6.0	2.6	6	2.4	5.9	2.4	5.9	2.4	5.9	2.4
		%	2.3	1.0	2.3	0.9	2.2	0.9	2.2	0.9	2.2	0.9
	RMSE	mm	3.5	1.4	3.5	1.3	3.4	1.3	3.4	1.3	3.4	1.3

4. Conclusions

Based on the proposed method assessment conducted in this study, several important conclusions can be drawn as follows:

- Uncertainty of vision-based tracking systems can be quantified by assessing the stereophotogrammetry results and conducting a noise floor test before running the real test. The image coordinate and reconstruction errors can inform about the degree of uncertainty of the stereophotogrammetry set. Further evaluation of the noise-floor tests is also strongly recommended as they do not only inform about the variance or bias errors in the monitoring system but also provide a preliminary insight into the uncertainty of the stereophotogrammetry sets in all possible
- Performing more than one set of stereophotogrammetry is recommended when accurate displacement measurements of a vision-based tracking system are needed. The uncertainty of stereophotogrammetry can still be a leading cause of displacement inaccuracy, which is

necessary to be checked in a simple way by evaluating either the image coordinates or the reconstruction errors.
– Using more images in stereophotogrammetry has proved insensitive to vision-based tracking system uncertainty and accuracy. For a similar FOV range as adopted for the monitored specimen in this study, a minimum of 10 images (5 images in each camera) is sufficient to obtain a high displacement accuracy. Note that the error presented is mostly valid for similar vision-based tracking settings and FOV as designed in this study.
– This study conclusively shows that systematic assessment of stereophotogrammetry and noise-floor tests should be part of routine vision-based tracking application in either static or dynamic monitoring. Comprehensive evaluations of stereophotogrammetry and noise-floor test results are strongly recommended not only as a preliminary check but also as a systematic approach to assess the high displacement measurement accuracy of aa vision-based tracking system.

Acknowledgments

The authors would like to thank the Los Alamos National Laboratory for providing funding for the vison-based system upgrade and validation tests in this study. We like to thank Sherif Elfass Ph.D. for valuable insight on the validation tests, and Tim Schmidt of Trilion Quantity Systems, USA for the technical support on various aspects of the system hardware and software.

References

[1] Martins, L. L., Rebordão, J. M., & Ribeiro, A. S. (2015). Structural observation of long-span suspension bridges for safety assessment: implementation of an optical displacement measurement system. In *Journal of Physics: Conference Series* (Vol. 588, No. 1, p. 012004). IOP Publishing. https://doi.org/10.1088/1742-6596/588/1/012004

[2] Brownjohn, J. M. W., Xu, Y., & Hester, D. (2017). Vision-based bridge deformation monitoring. *Frontiers in Built Environment*, *3*, 23. https://doi.org/10.3389/fbuil.2017.00023

[3] Poozesh, P., Sabato, A., Sarrafi, A., Niezrecki, C., & Avitabile, P. (2018). A multiple stereo-vision approach using three dimensional digital image correlation for utility-scale wind turbine blades. *Proc. IMAC XXXVI*, 12.

[4] Feng, D., & Feng, M. Q. (2017). Experimental validation of cost-effective vision-based structural health monitoring. *Mechanical Systems and Signal Processing*, *88*, 199-211. https://doi.org/10.1016/j.ymssp.2016.11.021

[5] Xu, Y., Brownjohn, J., & Kong, D. (2018). A non-contact vision-based system for multipoint displacement monitoring in a cable-stayed footbridge. *Structural Control and Health Monitoring*, *25*(5), e2155. https://doi.org/10.1002/stc.2155

[6] Ngeljaratan, L., & Moustafa, M. A. (2020). Structural health monitoring and seismic response assessment of bridge structures using target-tracking digital image correlation. *Engineering Structures*, *213*, 110551.

[7] Park, H. S., Park, J. S., & Oh, B. K. (2017). Vision-based stress estimation model for steel frame structures with rigid links. *Measurement Science and Technology*, *28*(7), 075104. https://doi.org/10.1088/1361-6501/aa6f50

[8] Jones, I., Iadicola, M. E. (Eds.). (2018). *A Good Practices Guide for Digital Image Correlation*. International Digital Image Correlation Society (IDICs). https://doi.org/10.32720/idics/gpg.ed1/print.format

[9] Ngeljaratan, L., Moustafa, M. A., & Pekcan, G. (2021). A compressive sensing method for processing and improving vision-based target-tracking signals for structural health monitoring. *Computer-Aided Civil and Infrastructure Engineering*. *36*(9), 1203-1223. https://doi.org/10.1111/mice.12653

[10] Ngeljaratan, L., & Moustafa, M. A. (2021). Underexposed Vision-Based Sensors' Image Enhancement for Feature Identification in Close-Range Photogrammetry and Structural Health Monitoring. *Applied Sciences*, *11*(23), 11086. https://doi.org/10.3390/app112311086

[11] Ngeljaratan, L., & Moustafa, M. A. (2020). Implementation and evaluation of vision-based sensor image compression for close-range photogrammetry and structural health monitoring. *Sensors*, *20*(23), 6844. https://doi.org/10.3390/s20236844

[12] Ngeljaratan, L., & Moustafa, M. A. (2019). System identification of large-scale bridges using target-tracking digital image correlation. *Frontiers in Built Environment*, *5*, 85. https://doi.org/10.3389/fbuil.2019.00085

[13] Olaszek, P. (1999). Investigation of the dynamic characteristic of bridge structures using a computer vision method. *Measurement*, *25*(3), 227–236. https://doi.org/10.1016/S0263-2241(99)00006-8

[14] Ribeiro, D., Calçada, R., Ferreira, J., & Martins, T. (2014). Non-contact measurement of the dynamic displacement of railway bridges using an advanced video-based system. *Engineering Structures*, *75*, 164–180. https://doi.org/10.1016%2Fj.engstruct.2014.04.051

[15] Zhu, L., Luo, H., & Zhang, X. (2009). Uncertainty and sensitivity analysis for camera calibration. *Industrial Robot: An International Journal*, *36*(3), 238–243. https://doi.org/10.1108/01439910910950496

[16] Jiang, T., Cui, H., Cheng, X., & Du, K. (2021). Calibration and uncertainty analysis of a combined tracking-based vision measurement system using Monte Carlo simulation. *Measurement Science and Technology*, *32*(9), 095007. https://doi.org/10.1088/1361-6501/abed85

[17] Caja, J., Gómez, E., & Maresca, P. (2015). Optical measuring equipments. Part I: Calibration model and uncertainty estimation. *Precision Engineering*, *40*, 298–304. https://doi.org/10.1016/j.precisioneng.2014.10.006

[18] Sims-Waterhouse, D., Isa, M., Piano, S., & Leach, R. (2020). Uncertainty model for a traceable stereo-photogrammetry system. *Precision Engineering, 63*, 1–9. https://doi.org/10.1016/j.precisioneng.2019.12.008

[19] Chaochuan, J., Ting, Y., Chuanjiang, W., Binghui, F., & Fugui, H. (2020). An extrinsic calibration method for multiple RGB-D cameras in a limited field of view. *Measurement Science and Technology*, *31*(4), 045901. https://doi.org/10.1088/1361-6501/ab48b3

[20] Chiodini, S., Pertile, M., Giubilato, R., Salvioli, F., Barrera, M., Franceschetti, P., & Debei, S. (2019). Experimental evaluation of a camera rig extrinsic calibration method based on retro-reflective markers detection. *Measurement*, *140*, 47–55. https://doi.org/10.1016/j.measurement.2019.03.036

[21] Reu, P. L. (2013). A study of the influence of calibration uncertainty on the global uncertainty for digital image correlation using a Monte Carlo approach. *Experimental Mechanics*, *53*(9), 1661–1680. https://doi.org/10.1007/s11340-013-9746-1

[22] Xia, R., Hu, M., Zhao, J., Chen, S., & Chen, Y. (2018). Global calibration of multi-cameras with non-overlapping fields of view based on photogrammetry and reconfigurable target. *Measurement Science and Technology*, *29*(6), 065005. https://doi.org/10.1088/1361-6501/aab028

[23] Wieneke, B. (2018). Improvements for volume self-calibration. *Measurement Science and Technology*, *29*(8), 084002. https://doi.org/10.1088/1361-6501/aacd45

[24] Discetti, S., & Astarita, T. (2014). The detrimental effect of increasing the number of cameras on self-calibration for tomographic PIV. *Measurement Science and Technology*, *25*(8), 084001. https://doi.org/10.1088/0957-0233/25/8/084001

[25] Lavatelli, A., & Zappa, E. (2016). Modeling uncertainty for a vision system applied to vibration measurements. *IEEE Transactions on Instrumentation and Measurement*, *65*(8), 1818–1826. https://doi.org/ 10.1109/TIM.2016.2541359

[26] Gu, F., Zhao, H., Ma, Y., & Bu, P. (2015). Camera calibration based on the back projection process. *Measurement Science and Technology*, *26*(12), 125004. https://doi.org/10.1088/0957-0233/26/12/125004

[27] Rodríguez-Quiñonez, J. C., Sergiyenko, O., Flores-Fuentes, W., Rivas-Lopez, M., Hernandez-Balbuena, D., Rascón, R., & Mercorelli, P. (2017). Improve a 3D distance measurement accuracy in stereo vision systems using optimization methods' approach. *Opto-Electronics Review*, *25*(1), 24–32. https://doi.org/10.1016/j.opelre.2017.03.001

[28] Cai, X., Zhong, K., Fu, Y., Chen, J., Liu, Y., & Huang, C. (2020). Calibration method for the rotating axis in panoramic 3D shape measurement based on a turntable. *Measurement Science and Technology*, *32*(3), 035004. https://doi.org/10.1088/1361-6501/abcb7e

[29] Niu, Z., Liu, K., Wang, Y., Huang, S., Deng, X., & Zhang, Z. (2017). Calibration method for the relative orientation between the rotation axis and a camera using constrained global optimization. *Measurement Science and Technology*, *28*(5), 055001. https://doi.org/10.1088/1361-6501/aa5fd4

[30] Liu, H., Zhou, X., Liu, Q., Ma, M., He, X., & Lin, J. (2021). Modeling, measurement, and calibration of three-axis integrated aerial camera pointing errors. *Measurement Science and Technology*, *32*(7), 075206. https://doi.org/10.1088/1361-6501/abdef5

[31] Xu, Y., Brownjohn, J. M., & Huseynov, F. (2019). Accurate deformation monitoring on bridge structures using a cost-effective sensing system combined with a camera and accelerometers: Case study. *Journal of Bridge Engineering*, *24*(1), 05018014. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001330

[32] Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(11), 1330-1334. https://doi.org/10.1109/34.888718

[33] Luo, P. F., Chao, Y. J., Sutton, M. A., & Peters, W. H. (1993). Accurate measurement of three-dimensional deformations in deformable and rigid bodies using computer vision. *Experimental Mechanics*, *33*(2), 123–132. https://doi.org/10.1007/BF02322488

[34] Zhu, C., Yu, S., Liu, C., Jiang, P., Shao, X., & He, X. (2019). Error estimation of 3D reconstruction in 3D digital image correlation. *Measurement Science and Technology*, *30*(2), 025204. https://doi.org/10.1088/1361-6501/aaf846

[35] Sutton, M. A., Yan, J. H., Tiwari, V., Schreier, H. W., & Orteu, J. J. (2008). The effect of out-of-plane motion on 2D and 3D digital image correlation measurements. *Optics and Lasers in Engineering*, *46*(10), 746–757. https://doi.org/10.1016/j.optlaseng.2008.05.005

[36] Schmidt, T., Tyson, J., & Galanulis, K. (2003). Full-field dynamic displacement and strain measurement-Specific examples using advanced 3D image correlation photogrammetry: Part II. *Experimental Techniques*, *27*(4), 22–26. https://doi.org/10.1111/j.1747-1567.2003.tb00118.x

[37] Schmidt, T., Tyson, J., & Galanulis, K. (2003). Full-field dynamic displacement and strain measurement using advanced 3D image correlation photogrammetry: Part I. *Experimental Techniques*, *27*(3), 47–50. https://doi.org/10.1111/j.1747-1567.2003.tb00115.x

Luna Nurdianti Ngeljaratan received her Bachelor degree in Universitas Gadjah Mada, Indonesia and earned her M.S degree from State University of New York at Buffalo, USA. She obtained her Ph.D. degree from the Department of Civil and Environmental Engineering, University of Nevada, Reno, USA, in 2019 where she worked as a Research Assistant and Post-Doctoral student of the co-author, Dr. Mohamed A. Moustafa. She is currently a researcher at the National Research and Innovation Agency of Indonesia with her research focused on structural health monitoring of civil infrastructures using vision-based systems, IoT sensor networks for structural damage indication and evaluation, and drone technology for rapid disaster mitigation.

Mohamed Aly Moustafa is an Associate Professor with the Civil and Environmental Engineering group at the University of Nevada, Reno. Dr. Moustafa received his M.S. and Ph.D. in Civil and Environmental Engineering from UC Berkeley. He also obtained a Certificate in Engineering and Business for Sustainability from UC Berkeley and is a registered professional engineer in the state of California. Moustafa's academic training and background is mainly in structural and earthquake engineering. He has more than 13 years' experience in mechanics and design of reinforced concrete structures, multi-scale experimental testing, structural and performance-based design for mitigating hazards, advanced monitoring techniques, and finite element and computational modeling with application to buildings, bridges, and infrastructure systems.