Next Generation 3D-DIC Technique with Sensor-Based Extrinsic Parameter Calibration and Natural Pattern Tracking

FABIO BOTTALICO, CHRISTOPHER NIEZRECKI
and ALESSANDRO SABATO

ABSTRACT

Three-dimensional Digital Image Correlation (3D-DIC) is a computer vision technique capable of extracting full-field displacements of structures by processing images acquired from synchronized stereo cameras and performing triangulation. To properly reconstruct the 3D points in space, the stereo cameras must be calibrated to compute the lens distortion (i.e., intrinsic parameters) and the cameras' relative position and orientation (i.e., extrinsic parameters). Traditionally, calibration is performed by taking pictures of a calibration target (e.g., a checkerboard). However, the calibration target must occupy most of the cameras' field of view (FOV) and have a size comparable to that of the object being tested, which becomes impractical for large-scale FOVs. Additionally, 3D-DIC requires the application of a stochastic speckle pattern on the surface to identify and track unique points across multiple images. This can be a problem when structures on which the application of a speckle pattern is impractical have to be analyzed (e.g., wind turbine blades). This research proposes a multi-sensor system using three inertial measurement units and a laser distance sensor to compute the extrinsic parameters of a stereo vision system and accelerate the calibration procedure. In addition, a feature-based natural pattern tracking algorithm is proposed to exploit features that are naturally present on the structure instead of applying a speckle pattern to perform 3D-DIC. Laboratory tests show that cameras calibrated using the multi-sensor system reconstruct displacements with an accuracy above 95% compared to displacements measured using a traditional calibration method. At the same time, the proposed feature-tracking algorithm can identify and reconstruct the 3D positions and displacements of preexisting features without requiring an applied speckle pattern. The research shows that the proposed methods can simplify and streamline the use of 3D-DIC and open the way to the use of this technique to perform condition monitoring of large-scale civil and mechanical engineering structures.

Fabio Bottalico, Christopher Niezrecki and Alessandro Sabato, Department of Mechanical and Industrial Engineering, University of Massachusetts Lowell, 1 University Avenue, Lowell, Massachusetts 01852, USA

INTRODUCTION

In recent years, computer vision techniques have been increasingly used for structural health monitoring (SHM) thanks to advancements in computational power and a decrease in the cameras' cost [1]. Compared to traditional contact-based methods, computer vision techniques provide full line-of-sight measurement capabilities where each pixel can be considered a sensor. Among the numerous computer vision techniques available, three-dimensional digital image correlation (3D-DIC) has been successfully used to extract displacements and strain profiles from images acquired from synchronized stereo cameras [2]. 3D-DIC has shown accuracy comparable to traditional contact-based methods and has been used to perform measurements on many structures [3-7].

Before performing 3D-DIC, **i)** the relative position and settings of the cameras must be known, and **ii)** a stochastic speckle pattern or optical targets need to be applied to the targeted structure. For the first point, a procedure called calibration is performed to determine the intrinsic and extrinsic parameters of the stereovision system [8]. Traditionally, calibration is performed by acquiring synchronized images of calibration objects such as a planar checkerboard. After recognizing the key features in the calibration object, the bundle adjustment algorithm is used to compute the camera parameters [9]. In order to perform a calibration, the size of the calibration object must be comparable to the size of the targeted structure. For structures larger than ~ 2 meters, custom-made calibration objects and calibration procedures must be used [5], which decreases the usability of 3D-DIC for large-scale SHM. In addition, once the intrinsic and extrinsic parameters of the stereovision system are calculated, any change in the cameras' relative position or settings will invalidate the calibration obtained. Therefore, cameras are typically fixed on a stiff bar to prevent relative motions, which limits the applicability of 3D-DIC. Therefore, more user-friendly and robust calibration procedures are researched that include collecting a single image of multi-planar calibration objects [10, 11], multiple images of a scale bar [12], or measuring the phase-shift of active targets [13, 14]. However, those procedures still suffer from all the drawbacks of traditional calibration, as they are image-based. Based on the separability of intrinsic and extrinsic parameters, researchers have recently proposed using auxiliary sensors to compute the extrinsic parameters. While some approaches are hybrid and use the sensor data to complement the collected images [15], a new approach to compute the extrinsic parameters using solely sensor data has recently been proposed [16-18].

In addition, to ensure that the same point of the targeted structure is individuated in the two synchronized images and image correlation can be performed, stochastic speckle patterns or optical targets are applied on the structure's surface to increase the robustness of the stereo match. However, it may not always be feasible for large-scale structures to apply a pattern/target, which further decreases the usability of 3D-DIC for large-scale SHM. For this reason, researchers started to develop methods to track naturally occurring patterns on the surface of the structure of interest. Recently, a method using KAZE features [19-21] has been proposed. To increase the applicability of 3D-DIC, this paper describes the experiments performed to characterize the accuracy of:

i) a novel multi-sensor system to compute the extrinsic parameters of a stereo vision system to streamline the calibration procedure and

ii) an alternative segmentation-based natural pattern tracking algorithm to eliminate the need for a patter/targets to perform 3D-DIC measurements.

Results of tests performed on lab-scale structures show that the sensor-based calibration yields an accuracy greater than 95% when compared to 3D-DIC measurements performed when a traditional image-based calibration is used. Also, the proposed natural pattern-tracking algorithm can reconstruct the 3D positions and displacements of natural features with average accuracy in X, Y, and Z of ~75%.

EXPERIMENTAL SETUP

Two tests were performed to evaluate the possibility of calibrating the stereo vision system using a multi-sensor board and performing 3D-DIC without a stochastic pattern. This section describes the experimental setups used to collect the data and validate the two proposed approaches. During Test #1, the multi-sensor board shown in Figure 1 developed by the authors in [17, 18] was used to calibrate a set of cameras and track the 3D displacement of a 0.8 x 0.6 m speckled test object moving in a ~4 x 3 x 3.5 m volume. The two cameras were placed at a distance of 4.4 m from each other and calibrated using traditional image-based and sensor-based calibrations. The traditional calibration was performed by collecting images of a checkerboard at different locations in the cameras' field of view (FOV) and using the bundle adjustment algorithm to compute the intrinsic and extrinsic parameters. The sensor-based calibration was performed by first calibrating the intrinsic parameters for each camera separately, using a checkerboard object, and then calibrating the extrinsic parameters using the data from the multi-sensor system. An overview of the experimental setup can be seen in Figure 2.

For Test #2, an impact test on an aluminum cantilever beam was performed (see Figure 3a). Random patterns were drawn on the beam using a marker to simulate naturally occurring stains or structural features. Immediately below one of the patterns, a traditional optical target was placed to allow the use of 3D-DIC for reference (see Figure 3b). Two Photron SA-2 high-speed cameras fitted with 24 mm lenses and placed at a baseline distance of 750 mm were used to capture the response of the beam with a sampling rate of 500 fps and extract the first three resonant frequencies of the structure. During the tests, the beam was impacted horizontally to ensure the main displacement occurred in the out-of-plane direction (i.e., Z).

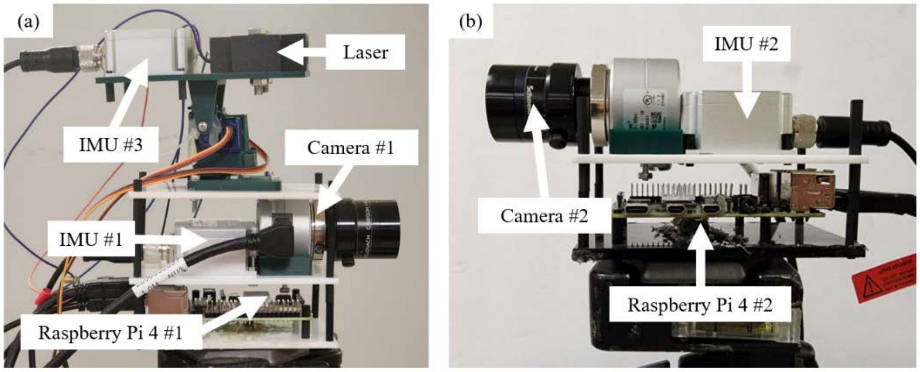


Figure 1. Overview of the prototype of the multi-sensor system: a) board #1 installed on the left camera of the stereo vision system and b) board #2 installed on the right camera.

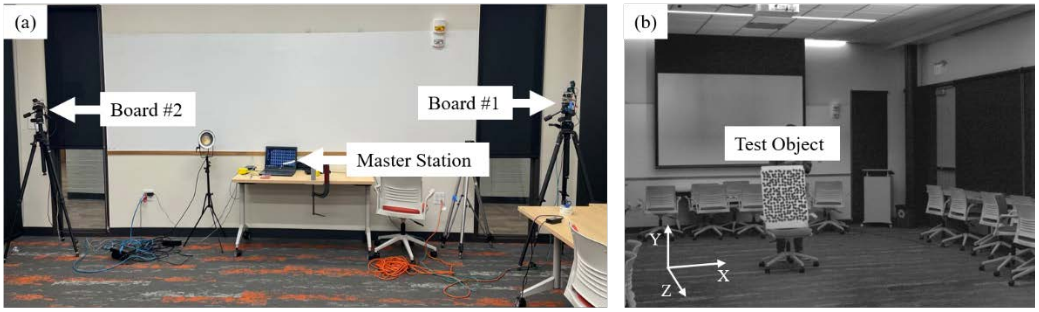


Figure 2. Experimental setup for Test #1: a) stereo vision system used to collect images of a test object and b) example of image captured by one of the cameras showing the speckled test object.

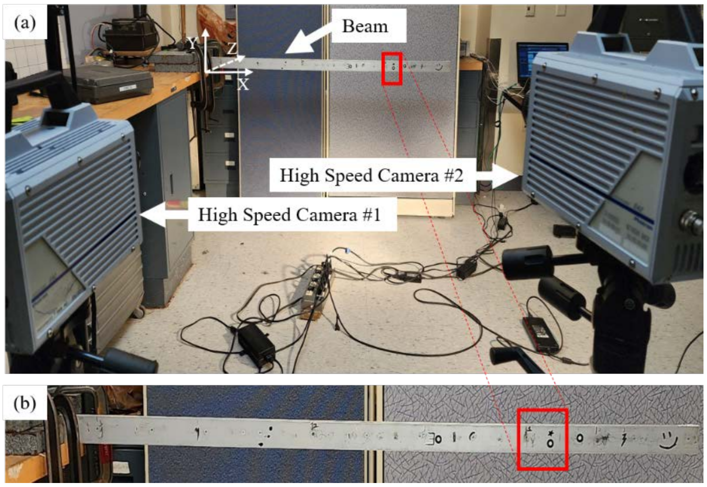


Figure 3. Experimental setup for Test #2: a) synchronized high-speed cameras used to collect a video of a vibrating cantilever beam and b) detail of the beam's surface showing the optical target and drawn shapes used to extract 3D displacements.

ANALYSIS OF THE RESULTS

In this section, the results of the tests described in the previous section are discussed with reference to the accuracy of the **i)** sensor-based calibration and **ii)** natural pattern tracking. In both scenarios, the collected images were processed using the open-source software Digital Image Correlation Engine (DICe) to extract the displacement of the targeted objects in the X, Y, and Z directions [22].

Results obtained with the cameras calibrated using the sensor-based method

The stereo images captured during Test #1 were processed twice using the calibration parameters obtained from the traditional image-based approach and the proposed sensor-based method. The goal of the test was to track the 3D displacement of the central point of the test object shown in Figure 2b and evaluate any difference in the results obtained when the two calibration procedures are used. Figure 4 shows the 3D trajectory of the test object and highlights any difference in the displacement extracted using the two methods. In particular, in Figure 4a, the color overlay represents the difference between the out-of-plane Z displacement computed with the traditional calibration and the one computed with the sensor-based calibration, while Figure 4b plots the difference between the in-plane X displacement computed with either calibration.

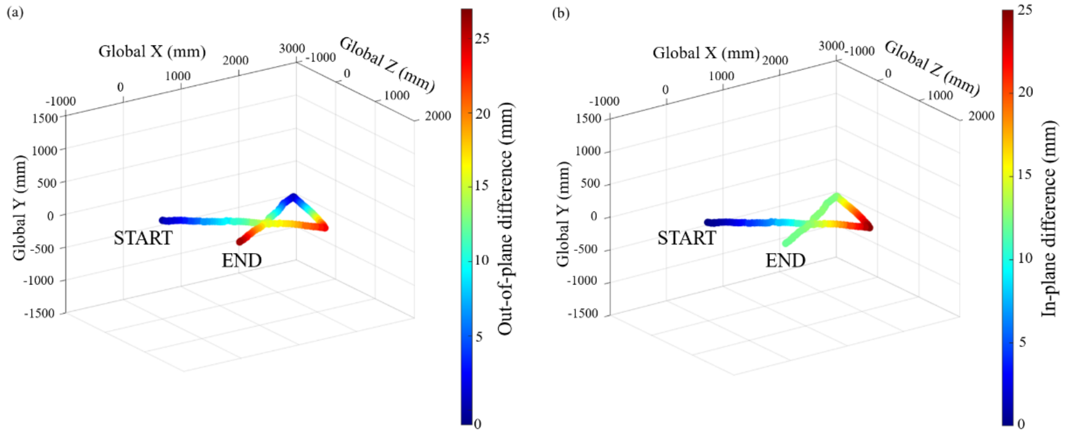


Figure 4. Trajectory of the test object measured with images processed using cameras calibrates with the image-based and the sensor-based calibrations: a) difference between the measured out-of-plane displacements, and b) the difference between the measured in-plane displacements.

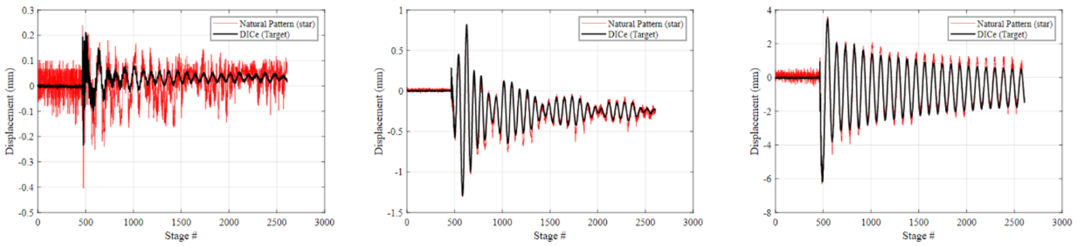
As observed from Figure 4a, the maximum out-of-plane displacement difference is equal to 27 mm over a Z-displacement of 1000 mm (i.e., relative error of 2.70%). Similarly, Figure 4b shows that the maximum in-plane displacement difference is equal to 25 mm over an X-displacement of 2200 mm (i.e., relative error of 1.14%), indicating an extremely high agreement between the two calibration methods and proving the validity of the sensor-based calibration approach.

Results obtained with the natural pattern tracking

For Test #2, the stereo images captured were processed using three different methods, namely DICe, the KAZE Strongest Interest Point (SIP) tracker described in [21], and the proposed centroid-based detector (CBD). In particular, DICe was used to track the displacement of the circular optical target highlighted in Figure 3b and used as ground truth. The proposed CBD and the literature available KAZE-SIP were used to track the displacement of the star-shaped pattern still visible in Figure 3b. The left and right displacements extracted with each camera were then triangulated using the calibration parameters computed before running the experiment and used to extract the 3D displacement in the global frame of reference shown in Figure 3a. Despite being one of the most commonly used algorithms to track features, KAZE-SIP has only been used to process images acquired from a single camera and without specifying any region of interest (ROI) [21]. Thus, all attempts to apply this method and perform matching between the two images collected with the stereo cameras failed after a few frames showing how this approach is not robust to perform 3D analyses.

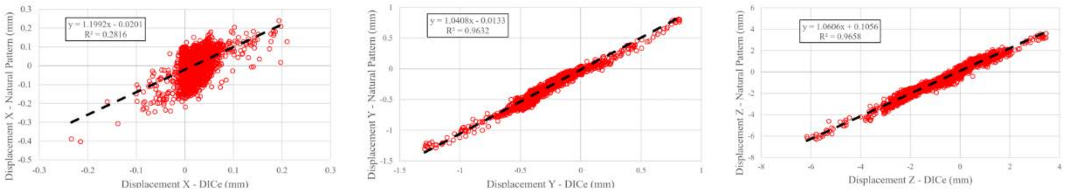
For the proposed CBD, an ROI is manually selected around the feature of interest in the first image (i.e., $t = t_0$) acquired with the left camera and used as starting point. The ROI is binarized to extract all the pixels belonging to the feature and calculate the centroid of the feature. Since the calibration parameters are known, using the fundamental matrix makes it possible to compute the epipolar line close to the feature's centroid in the first image acquired with the right camera. By selecting an appropriate search radius around the epipolar line, an ROI in the right image is identified and binarized to extract its centroid. The unique correspondence is found by matching the only feature in the left ROI with the multiple features in the right ROI. Once the centroids of the feature in the left and right images are found, a search window of arbitrary radius centered around the previously found centroids is generated for the following image (i.e., $t = t_1$). Under the assumption of small motions, such as in the case of acquisition at a high frame rate, it is possible to assume that the feature will not move by more than a couple of pixels between two consecutive times. Therefore, when the search windows are binarized, the feature is entirely inside the search windows and the new centroids are found. This cycle is then repeated for all the *n* frames.

The results of Test #2 are shown in Figure 5. The CBD is affected by significantly higher noise floors than the correlation-based tracker embedded in DICe. However, this is expected as feature-based methods have less subpixel resolution than correlation-based methods. Despite the higher noise content, the proposed CBD method calculates 3D displacement values that agree with those computed with DICe, and it can track displacements as small as 0.2 pixels. The Time Response Assurance Criterion (TRAC) of the two signals was computed to quantify the accuracy of the proposed CBD in measuring 3D displacement and verify its correlation with the data extracted with DICe. From the results shown in Figure 6, it is possible to observe how the correlation coefficient in the Z direction (i.e., the main direction of motion of the beam) is equal to 97%. Slightly less correlation is observed for the displacement in the Y direction (i.e., 96%), while for displacements close to the noise floor of the measurement, a correlation of 28% is obtained. While the results shown in this section are preliminary and need to be improved, it is possible to observe how the natural pattern tracker has the potential to become a valuable tool for performing dynamic measurements in environments where applying a target or a stochastic pattern is impractical. It should also be pointed out that in opposition to DICe, the proposed CBD method only requires the user to specify an ROI in the left image to initiate the process and that feature extraction, matching, and triangulation to obtain the 3D displacement is performed automatically.



(a) (b) (c)

Figure 5: Results from Test #2 showing the displacements of the beam computed with the three tracking methods: a) X-displacement, b) Y-displacement, and c) Z-displacement.



(a) (b) (c)

Figure 6. Correlation analysis between the displacements computed with DICe (i.e., reference) and the proposed CBD natural pattern tracker method: a) X-displacement, b) Y-displacement, and c) Z-displacement.

CONCLUSIONS

This study proposes and investigates two novel methods to measure the extrinsic parameters of a stereovision system for 3D-DIC and for tracking the natural pattern of a structure. By measuring the relative orientation and position of two cameras using three IMUs and a laser, the proposed multi-sensor method can compute the extrinsic parameters of the stereovision system. When measuring the 3D displacements of a planar object moving in a large-FOV test, measurements performed with the sensor-based calibration differ from measurements performed with a traditional image-based calibration by ~1.1% in the in-plane direction and ~2.7% in the out-of-plane direction. Additionally, by measuring the location of the centroid of a feature and using epipolar geometry, it is possible to automatically detect, match, track, and triangulate the location of a naturally occurring feature on the surface of a structure of interest. When comparing the measured 3D displacements with those obtained using the open-source software DICe, an average TRAC value equal to 75% is obtained, showing an excellent agreement between the natural pattern methods and correlation-based tracking, despite a higher sensitivity to noise. The results shown in this research prove how the proposed methods can be valid alternatives to traditional image-based calibration and speckle pattern tracking and, if further improved, can increase the applicability of 3D-DIC for monitoring large-scale structures.

ACKNOWLEDGMENT

This work was supported by the U.S. National Science Foundation (NSF) under award number 2018992, "*MRI: Development of a calibration system for stereophotogrammetry to enable large-scale measurement and monitoring*." The contents are those of the authors and do not necessarily represent the official views of the funding agency.

REFERENCES

1. D. Feng, and M. Q. Feng, 2018. "Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection – A review" in *Engineering Structures*, 156, pp. 105-117.
2. M. A. Sutton, J.J. Orteu, and H. W. Schreier, 2009. "Two-Dimensional and Three-Dimensional Computer Vision" in *Image Correlation for Shape, Motion and Deformation Measurements: Basic Concepts, Theory and Applications.* Boston: Springer, pp. 1-16.
3. A. Sabato, and C. Niezrecki, 2017. "Feasibility of Digital Image Correlation for railroad tie inspection and ballast support assessment" in *Measurement*, 103, 93-105.
4. L. Ngeljaratan, and M. A. Moustafa, 2020. "Structural health monitoring and seismic response assessment of bridge structures using target-tracking digital image correlation" in *Engineering Structures*, 213, 110551.
5. P. Poozesh, A. Sabato, A. Sarrafi, C. Niezrecki, P. Avitabile, and R. Yarala, 2020. "Multicamera measurement system to evaluate the dynamic response of utility-scale wind turbine blades" in *Wind Energy*, 23(7), 1619-1639.
6. D. Reagan, A. Sabato, and C. Niezrecki, 2018. "Feasibility of using digital image correlation for unmanned aerial vehicle structural health monitoring of bridges" in *Structural Health Monitoring*, 17(5), 1056-1072.

7. M. Kalaitzakis, N. Vitzilaios, D. C. Rizos, and M. A. Sutton, 2021. “Drone-Based StereoDIC: System Development, Experimental Validation and Infrastructure Application” in *Experimental Mechanics*, 61(6), 981-996.
8. Z. Zhang, 2014. “Camera Parameters (Intrinsic, Extrinsic)” in *Computer Vision: A Reference Guide*, K. Ikeuchi ed. Boston: Springer pp. 81-85.
9. B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, 2000. "Bundle Adjustment — A Modern Synthesis" in *Vision Algorithms: Theory and Practice*, B. Triggs, A. Zisserman, R. Szeliski eds. Berlin, Heidelberg: Springer, pp. 298-372.
10. J. Zhang, J. Zhu, H. Deng, Z. Chai, M. Ma, and X. Zhong, 2019. “Multi-camera calibration method based on a multi-plane stereo target” in *Applied Optics*, 58(34), 9353-9359.
11. D. Solav, K. M. Moerman, A. M. Jaeger, K. Genovese, and H.M. Herr, 2018. “MultiDIC: An open-source toolbox for multi-view 3D digital image correlation” in *IEEE Access*, 6, 30520-30535.
12. P. Sun, N.-G. Lu, M.-L. Dong, B.-X. Yan, and J. Wang, 2018. “Simultaneous All-Parameters Calibration and Assessment of a Stereo Camera Pair Using a Scale Bar” in *Sensors*, 18(11), 3964.
13. B. Chen, K. Genovese, and B. Pan, 2022. “Calibrating large-FOV stereo digital image correlation system using phase targets and epipolar geometry” in *Optics and Lasers in Engineering*, 150, 106854.
14. K. Genovese, Y. Chi, and B. Pan, 2019. “Stereo-camera calibration for large-scale DIC measurements with active phase targets and planar mirrors” in *Optics Express*, 27(6), 9040-9053.
15. W. Feng, Z. Su, Y. Han, H. Liu, Q. Yu, S. Liu, D. Zhang, 2020. “Inertial measurement unit aided extrinsic parameters calibration for stereo vision systems” in *Optics and Lasers in Engineering*, 134, 106252.
16. A. Sabato, N. A. Valente, and C. Niezrecki, 2020. “Development of a Camera Localization System for Three-Dimensional Digital Image Correlation Camera Triangulation” in *IEEE Sensors Journal*, 20(19), 11518-11526.
17. F. Bottalico, C. Niezrecki, K. Jerath, Y. Luo and A. Sabato, 2023. “Sensor-Based Calibration of Camera’s Extrinsic Parameters for Stereophotogrammetry” in *IEEE Sensors Journal*, 23(7), 7776-7785.
18. F. Bottalico, N. A. Valente, C. Niezrecki, K. Jerath, Y. Luo, and A. Sabato, 2023. “Sensor-aided Camera Calibration for Three Dimensional Digital Image Correlation Measurements” in *Proc. SPIE* 12488.
19. D. Kumar, C.-H. Chiang, and Y.-C. Lin, 2022. “Experimental vibration analysis of large structures using 3D DIC technique with a novel calibration method” in *Journal of Civil Structural Health Monitoring*, 12(2), 391-409.
20. D. Kumar, C.-H. Chiang, and S. Prasad, 2022. “Integrating robust feature detection methodology with in-house DIC for identification and correlation of natural patterns on large structures” in *Proc. SPIE* 12047.
21. S. Prasad, C.-H. Chiang, D. Kumar, S. Kalra, and A. Khandelwal, 2023. “Robust and efficient feature-based method for structural health monitoring of large structures”, in *Journal of Civil Structural Health Monitoring.*
22. D. Z. Turner, 2015. “Digital Image Correlation Engine (DICe) Reference Manual” Sandia Report SAND2015-10606 O.