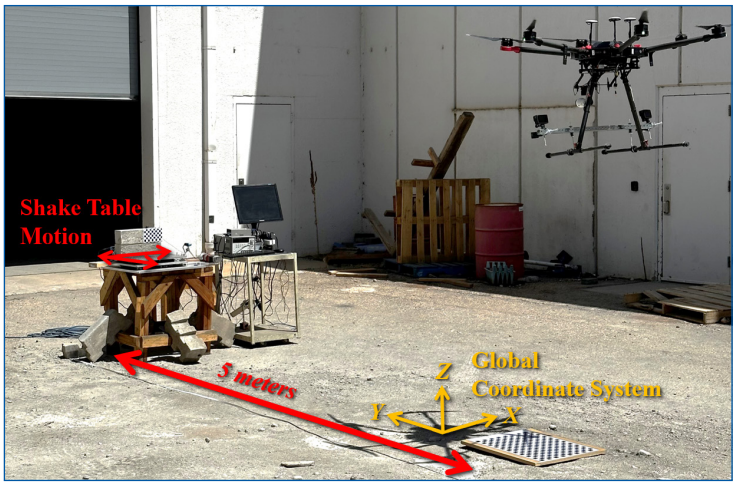
MOUNTAIN-PLAINS CONSORTIUM

MPC 24-553 | B.J. Perry, Y. Guo, and R. Atadero

NONCONTACT DYNAMIC THREE-COMPONENT DISPLACEMENT MEASUREMENT WITH A DUAL STEREOVISION-ENABLED UNCREWED AERIAL SYSTEM



MOUNTAIN-PLAINS CONSORTIUM REGION 8

A University Transportation Center sponsored by the U.S. Department of Transportation serving the Mountain-Plains Region. Consortium members:

Colorado State University
North Dakota State University
South Dakota State University
University of Colorado Denver
University of Denver
University of Utah
Utah State University
University of Wyoming

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
MPC-680		

4. Title and Subtitle	5. Report Date
Noncontact Dynamic Three-Component Displacement Measurement with a Dual Stereovision-Enabled Uncrewed Aerial System	September 2024
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Brandon J. Perry Yanlin Guo Rebecca Atadero	MPC 24-553

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Colorado State University Department of Civil and Environmental Engineering Fort Collins, CO 80523	11. Contract or Grant No.

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Mountain-Plains Consortium North Dakota State University PO Box 6050, Fargo, ND 58108	Final Report
	14. Sponsoring Agency Code

15. Supplementary Notes

Supported by a grant from the US DOT, University Transportation Centers Program

16. Abstract

Measuring the dynamic displacements of a structure provides a comprehensive understanding of the structure, especially when subjected to different types of dynamic loading (i.e., wind, traffic, impact loads, blast loads). Despite their usefulness, direct displacement measurements are typically not collected due to the cumbersome logistical issues of sensor placement and maintenance and the impracticality of instrumenting contact-based sensors across all significant structures. In this context, this study proposes a novel dual stereo vision technique to measure the dynamic displacement of structures using a portable, non-contact measurement system that involves an uncrewed aerial system (UAS) and four optical cameras. One pair of cameras tracks the three-component (*x*, *y*, and *z*) motion of a region of interest (ROI) on a structure with respect to the UAS system; the other pair of cameras measures the six degrees of freedom motion (6-DOF) (both rotational and translational motion) of the UAS system by tracking a stationary reference. The motion of the UAS is then compensated for to recover the true dynamic displacement of the ROI. The proposed dual stereo vision technique realizes simultaneous measurement of all three components of the structure's displacements and 6-DOF of UAS motion through a mathematically elegant process. The unique dual stereo vision technique allows flexibility in choosing a global reference coordinate system, greatly enhancing the feasibility of applying the new technology in various field environments. This new technique has overcome the major challenge of significant UAS motions in full-scale applications. Furthermore, this technique relies on natural features and eliminates the requirement of artificial targets on the structure, permitting applications to difficult-to-access structures.

17. Key Word	18. Distribution Statement
bridges, decision making, inspection, risk management	Public distribution

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	36	n/a

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Noncontact Dynamic Three-Component Displacement Measurement with a Dual Stereovision-Enabled Uncrewed Aerial System

Brandon J. Perry, S.M.ASCE[1]
Yanlin Guo, Ph.D., A.M.ASCE[2]
Rebecca Atadero, Ph.D., M.ASCE[3]

[1]Dept. of Civil and Environmental Engineering, Colorado State Univ., Fort Collins, CO. Email: bjperry@colostate.edu
[2]Dept. of Civil and Environmental Engineering, Colorado State Univ., Fort Collins, CO (corresponding author). ORCID: https://orcid.org/0000-0002-7162-6508. Email: yanlin@colostate.edu
[3]Dept. of Civil and Environmental Engineering, Colorado State Univ., Fort Collins, CO. Email: Rebecca.atadero@colostate.edu

September 2024

Acknowledgments

This project is funded by Mountain-Plains Consortium (MPC) under grant number MPC-680. The opinions and findings of the authors do not necessarily reflect the view and opinions of MPC.

The work presented in this paper was conducted with support from Colorado State University (CSU) and the Mountain-Plains Consortium, a University Transportation Center funded by the U.S. Department of Transportation (FASTACT Grant No. 69A3551747108). The contents of this paper reflect the views of the authors, who are responsible for the facts and accuracy of the information presented. Additionally, the authors would like to acknowledge the Colorado State University's Drone Center and Mr. Christopher Robertson for providing the UAVs and expertise, as well as Mr. Todd Atadero of Colorado State University for his insight and assistance on the experiments in this study.

Data Availability Statement

The camera data used in the stationary point measurement experiment are available from the corresponding author upon reasonable request.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

ABSTRACT [1]

Measuring the dynamic displacements of a structure provides a comprehensive understanding of the structure, especially when subjected to different types of dynamic loading (i.e., wind, traffic, impact loads, blast loads). Despite their usefulness, direct displacement measurements are typically not collected due to the cumbersome logistical issues of sensor placement and maintenance and the impracticality of instrumenting contact-based sensors across all significant structures. In this context, this study proposes a novel dual stereo vision technique to measure the dynamic displacement of structures using a portable, non-contact measurement system that involves an uncrewed aerial system (UAS) and four optical cameras. One pair of cameras tracks the three-component (x, y, and z) motion of a region of interest (ROI) on a structure with respect to the UAS system; the other pair of cameras measures the six degrees of freedom motion (6-DOF) (both rotational and translational motion) of the UAS system by tracking a stationary reference. The motion of the UAS is then compensated for to recover the true dynamic displacement of the ROI. The proposed dual stereo vision technique realizes simultaneous measurement of all three components of the structure's displacements and 6-DOF of UAS motion through a mathematically elegant process. The unique dual stereo vision technique allows flexibility in choosing a global reference coordinate system, greatly enhancing the feasibility of applying the new technology in various field environments. This new technique has overcome the major challenge of significant UAS motions in full-scale applications. Furthermore, this technique relies on natural features and eliminates the requirement of artificial targets on the structure, permitting applications to difficult-to-access structures.

[1] This report is based on the contents from the following published journal paper:
Perry, B.J., Guo, Y., and Atadero, R. (2024), "Noncontact Dynamic Three-Component Displacement Measurement with a Dual Stereovision-Enabled Uncrewed Aerial System," *ASCE*, Vol. 150, No. 8, 04024052.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

This report introduces a novel method for measuring the dynamic displacements of structures using a dual stereovision-enabled uncrewed aerial system (UAS). Traditional methods for displacement measurement often face challenges due to logistical issues of sensor placement and maintenance. This study addresses these challenges by utilizing a portable, non-contact system involving a UAS equipped with four optical cameras.

Key Findings:

1. **Dual Stereovision Technique**: The study presents a dual stereovision technique that uses two pairs of cameras. One pair tracks the three-component (3C) motion (x, y, and z) of a region of interest (ROI) on a structure relative to the UAS, while the other pair measures the six degrees of freedom (6-DOF) motion of the UAS by tracking a stationary reference.
2. **Calibration**: A unique dual stereo vision calibration technique was developed to overcome the challenge of calibrating four cameras facing different directions. This technique ensures precise measurement of the 3C displacements and 6-DOF UAS motion.
3. **Experimental Validation**: The proposed system was validated through laboratory experiments and UAS-based measurements. The system demonstrated high accuracy in measuring the dynamic displacements of structures, with sub-millimeter RMS error in stationary tests and a few millimeters RMS error in dynamic tests.
4. **Applications and Feasibility**: The system can be applied to various field environments, including difficult-to-access structures, without the need for artificial targets. The flexibility in choosing a global reference coordinate system enhances its practical applicability.

Conclusions and Future Work: The dual stereovision technique successfully measures the dynamic displacements of structures, addressing major challenges in using UAS for such measurements. Future improvements could include enhancing camera synchronization and exploring methods to reduce ground sampling distance (GSD) for better accuracy at greater distances. The study demonstrates the potential for widespread application of this technology in structural health monitoring and other civil engineering fields.

1. INTRODUCTION

Measurements of dynamic displacements of a structure allow for a holistic understanding of structural behaviors under various dynamic loading, such as traffic, wind, impact loading, and blast loading. The measured dynamic displacements can be used to identify dynamic properties of the structure [1], [2], [3], [4], [5], [6], [7], abnormal patterns of structural response [8], and/or excessive loading and are, therefore, valuable for assessing the health of the structure. Traditionally, displacements are measured through contact-based sensors such as accelerometers (indirectly through double integration), linear variable differential transformers (LVDTs) [9], [10], or global-positioning systems (GPSs) attached to a structure. Logistical issues such as installation and maintenance costs may impede the implementation of contact-based sensors [11], [12]. If a wireless system is installed, time synchronization and communication loss become additional issues [12]. Moreover, once a sensor suite is installed on the structure, it is often difficult to move the sensors to another location on the structure or to move the sensors onto a different structure. It is also impracticable to instrument all significant structures with contact-based sensors. Even with the widely used accelerometers, they typically have limitations in capturing low-frequency content in displacements [13]. As an alternative, stationary contact-free remote sensors, such as cameras or laser Doppler, have been deployed to measure displacements [14], [15], [16], [17], [18], [19], [20], [21], [22], [23]. However, finding a safe placement area for the stationary system may not be feasible in some applications.

As a portable alternative, an uncrewed aerial system (UAS; also referred to as drones)-enabled remote sensing has been explored to measure displacement for structures in difficult-to-access locations. UAS-enabled sensors have been used to track the displacements of a structure in either the two-component (2C) planar directions (the plane that is perpendicular to the camera) or the one-component (1C) depth direction (the out-of-plane direction/the distance from the camera to the structure). The major challenge of using a UAS-based sensing system is that while the UAS can hover precisely in the air, there is still a level of its drift and rotation, which may introduce measurement errors. Ideally, the motion of the UAS needs to be measured and accounted for in implementation. Although the trajectory of the UAS can be measured using GPS and RTK technologies, the precision of the UAS position (*cm* level) with these technologies is not precise enough for dynamic displacement measurements. Some existing studies assumed that UAS drift is minimal and did not compensate for the errors due to UAS motion. For example, Catt et al. implemented passive stereo vision to measure the out-of-plane deformations of a deformable board with a random speckle pattern [24]. Yoon et al. [25] used a UAS platform to identify the vibration modes of a structure. Other studies have explored correcting the errors induced by UAS motion. For instance, one study used a high-pass filter to pre-process the data, assuming that the random drifts occur at a lower frequency than the structure being studied [26], [27]. Another technique is to use a stationary reference point. Kalaitzakis et al. implemented digital image correlation to measure the strain during a four-point loading test with a painted speckle pattern on both a concrete beam and a stationary reference in the camera's foreground [28]. Garg et al. used accelerometers attached to the UAS to measure the yaw and pitch of the sensor (laser Doppler vibrometer in this case). Although some site-sensor or application-specific techniques were able to treat the UAS drift issue to some extent (i.e., [28], [29]), fully compensating for the movement of the UAS in a more general approach is still an ongoing challenge for UAS-enabled dynamic displacement measurements.

In addition to tracking the motion of the UAS, there are still additional limitations to overcome to make UAS-enabled dynamic displacement measurement viable. Many of the previously mentioned camera-based techniques implement an artificial pattern or target to provide accurate measurements (such as [24], [26], [28], [30], [31], [32]); however, such a technique would prove problematic in the field when accessibility is an issue. Moreover, only taking measurements within an artificial target area would not take advantage of the camera's entire field of view (FOV). Using only natural features inherent in a structure or the surroundings would provide greater system feasibility and ease of use for an end user.

Yoon et al. proposed a 2C planar UAS-enabled measurement methodology without a speckle pattern by identifying, matching, and tracking keypoints in the background of the video; however, the proof of concept still required LED lights as targets [33]. Lastly, in all the previously mentioned studies of UAS-enabled dynamic displacement measurements, the studies focused on measuring only 1C (out-of-plane depth) or 2C (planar directions perpendicular to the camera). The existing studies have yet to attempt to measure the full 3C displacements of a structure using UAS-enabled sensors without the aid of a painted speckle pattern, which can be critical for certain types of structures such as long-span bridges and cables. Perry and Guo [30] proposed a new technique that measures all 3C displacements, where a UAS was instrumented with integrated optical and IR sensors. In this technique, the video of 2C planar movement is captured by the optical sensor and, simultaneously, the 1C depth movement is recorded by the IR sensors in conjunction with an IR projector using active stereo vision. However, the concept in Perry and Guo [30] has only been proven in a small-scale laboratory setting with a close camera-to-structure distance.

Considering the advantages of UAS technology, lessons learned, and limitations of the previous studies, this paper proposes to scale up the concept of 3C displacement measurement first proposed in Perry and Guo [30] for application to full-scale structures, where the stereo vision technique is utilized to measure the 3C displacement of the structure and compensate for the motion of the UAS by tracking a stationary reference target. Despite the straightforward concept, significant challenges are associated when scaling up this technique for large-scale applications.

Challenge #1: Calibration of four cameras in dual stereo vision pairs. In large-scale applications, it is difficult to capture both the region of interest (ROI) on structures and the reference target in the same image due to the 3D layout of the structure, which necessitates the use of two pairs of cameras with one pair tracking the ROI and the other pair tracking the reference target. Although camera calibration (i.e., solving internal camera parameters) and stereo vision rectification (i.e., solving external camera parameters between a camera pair) have been well studied, calibrating four cameras together for dual stereo vision measurement is very challenging since the four cameras facing different directions cannot capture the same calibration targets. This type of dual stereo vision calibration has not been reported in the literature. This study proposes a unique dual stereo vision calibration technique to tackle this challenge.

Challenge #2: Precisely accounting for the six degrees-of-freedom (6-DOF) motion of a UAS. Large-scale applications require a longer working distance from the cameras to a structure for safety and accessibility considerations (the working distance is defined in Figure 1.1 and Figure 1.2). To achieve high measurement accuracy at a further distance, a longer baseline (Figure 1.1 and Figure 1.2) between two cameras in each pair is needed. With long camera baselines, the geometric measurement errors due to the three degrees-of-freedom (3-DOF) rotation of the UAS system will be magnified significantly. Thus, compensating for the rotation of the UAS becomes a critical yet challenging step due to the multiple DOFs involved. Although some earlier studies have explored compensating for the translational motions of UAS [28], [29], [30], there have been no existing methodologies to compensate the 3-DOF rotation of UAS for 3C dynamic displacement measurements using computer vision in the literature. To overcome this challenge, this study proposes a mathematically elegant method to simultaneously measure the 3-DOF rotation and 3-DOF translational motions of the UAS.

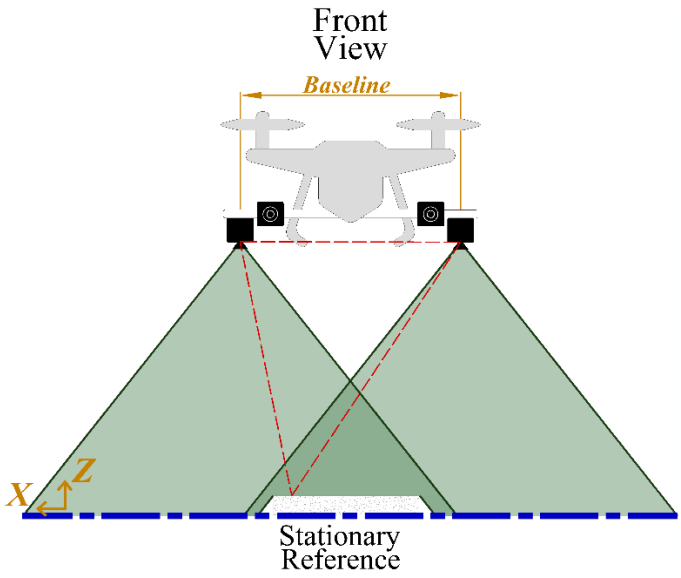


Figure 1.1 Overview of dual stereovision camera setup: front view of the UAS system with the overlap of the stereo cameras shown

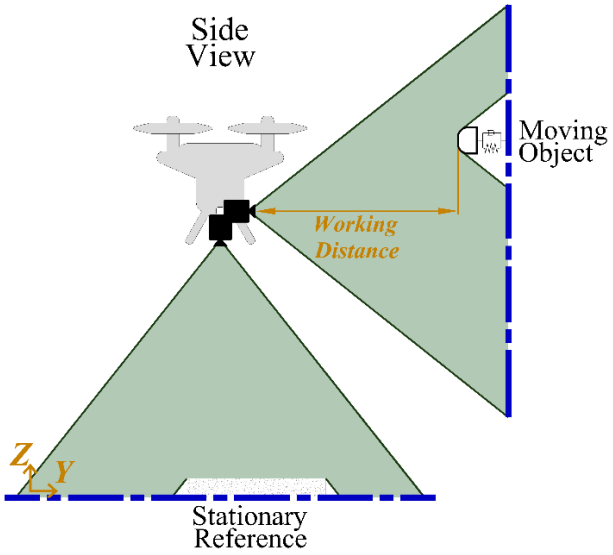


Figure 1.2 Overview of dual stereovision camera setup: side view of the UAS system with the layout of the dual stereovision system

In addition to the above two challenges, tracking the natural features of structures is another important consideration for practice. Attaching artificial patterns or targets to a structure might be impractical in practice [24], [28], [30]. To the authors' knowledge, there is no successful implementation using only natural features to directly measure the 3C displacements of a structure with a non-stationary camera. Although there is a slight tradeoff between measurement accuracy and ease of implementation, this study will explore how to track natural features in concrete with minimal user input.

In the proposed technique, four total optical sensors are used: two cameras pointed at the ROI to record the dynamic displacement of the structure and two additional cameras pointed at a stationary reference to track the movement of the UAS. Each camera pair implements passive stereo vision to record the dynamic displacements without the need for an artificial target or laser projector. The 6-DOF

displacements of the UAS are measured and compensated for in the ROI measurements. This dual passive stereo vision framework is novel and provides a portable system for structural surveillance compared with stationary non-contact sensors while still being able to measure the 3C dynamic displacement using the full FOV of the camera.

In this paper, the methodology section presents techniques to address the two major challenges. Next, experimental results are shown to evaluate the effectiveness of the methodologies. The first experiment is specifically designed to stress-test the two challenges, while the second experiment explores the practicality of the system with the cameras attached to a UAS. Lastly, there is a discussion of the advantages, limitations, and lessons learned from the system, with a conclusion summarizing the merits of the proposed technique.

2. METHODOLOGY

This section will discuss the methodology of the dual stereo vision system. First, the design of the camera rig, which holds the four cameras, and the UAS used will be discussed. Next, the methodology to calibrate the four cameras to find the relative location of the cameras is described. Lastly, the proposed technique to measure the displacement of the ROI and to compensate for the UAS movement is introduced.

2.1 UAS-Based Camera Platform

Four GoPro Hero 10s (hereafter, GPs) were chosen because of their compact size, sufficient battery, and high-quality imagery. The GPs had a superior ground sampling distance (GSD) (the distance between two consecutive pixel centers measured on the real-world target) due to a greater sensor size and number of captured pixels than those used in the previous study [30]. From a 5-m distance, the GPs had a GSD of 1.99 mm/pixel. Additionally, from a 5-m distance, the cameras captured an area of 5.9 m × 10.6 m. With a 1/2.3-in. sensor (6.17 mm × 4.55 mm), the GPs were able to capture 16 MP images at a sampling rate of 60 frames/sec (FPS). The GPs were attached to a DJI Matrice 600 Pro UAS. This UAS can carry a 5.5 kg payload while providing about 20 minutes of flight time.

The two camera pairs were placed at a 90° angle from each other, with one pair tracking a moving object (structure) and the other pair tracking a stationary target, as shown in Figure 1.1 and Figure 1.2. Note that the stationary reference is not required to be under the UAS; likewise, the ROI is not required to be in front of the UAS, as shown in Figure 1.1 and Figure 1.2. Rather, the stereo camera pairs can be placed at any angle from each other according to the geometries of the structure of interest and available reference in field applications. The configurations of the four GPs and the UAS are shown in Figure 2.1 and Figure 2.2. The entire camera rig was designed to be as rigid as possible to ensure that the four cameras moved in tandem (i.e., there is no relative motion among cameras), which is vital for accurate measurements. Each camera was bolted to a square steel tube, and the bolted connections were secured with a permanent thread lock. The rigid steel tube restricted the movement between the GPs (i.e., consistent baseline). Next, to reduce the camera's vibration while attached to the UAS, rubber washers, rubber grommets, and rubber rings were attached to the connections between the UAS and the square steel tube. The steel tubing and bolted cameras are referred to as the *camera rig*. The camera rig was removed from the UAS during calibration and lab testing for easier handling. The gimbaled camera originally attached to the Matrice 600 was still operable during the flight for easier navigation and control.

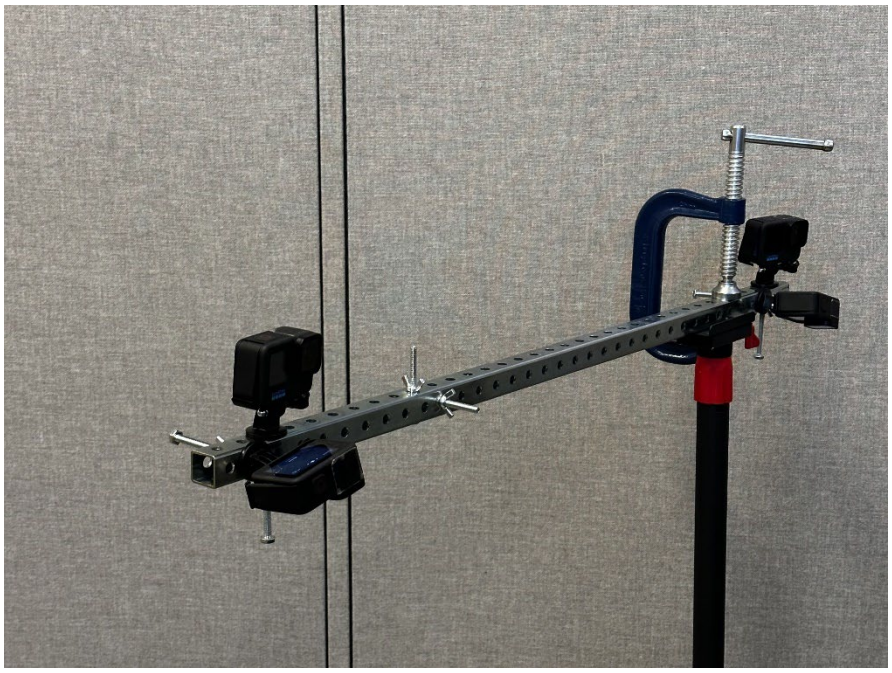


Figure 2.1 Dual stereovision camera rig attached to a DJI Matrice 600 Pro UAS: detached camera rig used during calibration and rectification

Figure 2.2 Dual stereovision camera rig attached to a DJI Matrice 600 Pro UAS:Matrice 600 Pro with attached camera rig

2.2 Dual-Pair Camera Calibration

Before taking measurements, the cameras need to be calibrated. The dual stereo vision process requires a three-step calibration process:

1. Individual camera calibration to undistort images
2. Stereo rectification to solve for the parameters of transformation (extrinsic) matrix of a camera pair and virtual intrinsic parameters
3. Dual stereo pairs calibration to find the location of the camera pairs with respect to each other and to unify measurements to one local coordinate system

Step 1: Individual camera calibration. The dual stereo pairs calibration is a novel technique, while individual camera calibration and rectification have been previously well researched. Typical camera calibration identifies the *Intrinsic Matrix*, **I**, [Eq. (1)] and distortion coefficients, which remove distortions in the image [34]. **I** consists of the camera's focal lengths in the *x*- and *y*- directions (f_x and f_y, respectively), the skew factor, s, and the principal points of the image (c_x and c_y). In this implementation, the skew factor, s, was assumed to be 0, a typical assumption for modern cameras. The camera parameters were found using MATLAB's built-in Camera Calibration App with a checkerboard pattern. A high-quality checkerboard pattern was printed on thick paper and glued to a sheet of glass to ensure a smooth, flat surface, which is important for accurate calibration and rectification. The results were then exported for later use. Note that the camera calibration must be performed on every camera individually (i.e., four total sets of calibration parameters); however, the parameters remain constant for each individual camera unless there is a significant knock to the camera or a large temperature shift [35].

$$[\mathbf{I}]_{3\times4} = \left[\begin{array}{ccc|c} f_x & s & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \quad (1)$$

Step 2: Stereo vision rectification. Each camera pair must be rectified once the cameras are calibrated and the distortions removed from the images. Rectification ensures that epipolar lines exist between the two cameras (i.e., the vertical pixel location of the *ROI* is the same between the two cameras) and that the image planes of the camera are on the same plane. Stereo rectification solves for the rotation, $\mathbf{R}_{r\to l}$, and translation, $\mathbf{T}_{r\to l}$, of the second camera (i.e., right camera in this implementation) with respect to the first camera (i.e., left camera). The rotation matrix and translation vector can then be combined to build a transformation matrix (also known as extrinsic matrix), [$\mathbf{A}_{r\to l}$] [Eq. (2)], which transforms the local coordinate system of the right camera to that of the left camera of the camera pair. This process was performed with MATLAB's built-in Stereo Camera Calibrator App with a checkerboard pattern glued to a glass plane. After rectification, the two image planes are assumed to be on the same plane. This transformation changes the focal length, principal point, and baseline found in the camera calibration procedures. Therefore, a virtual focal length, principal point, and baseline are estimated based on the rectification parameters. These virtual parameters are used in subsequent steps to find the 3D distances. The camera pair rectification must be performed for each camera pair (i.e., two times total) but does not need to be re-performed unless the cameras are moved or rotated with respect to each other from their original position. During the camera calibration and rectification procedures, the camera rig was suspended above the ground with a tripod with the calibration plane placed on the ground. The cameras were triggered remotely using a proprietary remote for the GPs. This remote triggered the GPs to start capturing videos at roughly the same time with an error within a couple of frames. Although this synchronization error is too large for measuring dynamic displacement, it is not an issue for calibration and rectification since the cameras and calibration board are in a stationary position for each video capture.

$$[\mathbf{A}_{r\to l}]_{4\times4} = \begin{bmatrix} \mathbf{R}_{r\to l} & \mathbf{T}_{r\to l} \\ 0 & 1 \end{bmatrix}_{4\times4} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & T_1 \\ R_{21} & R_{22} & R_{23} & T_2 \\ R_{31} & R_{32} & R_{33} & T_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Step 3: Dual stereo pairs calibration. The dual pairs calibration is performed to measure the position of each camera pair with respect to the other camera pair. A novel dual stereo pairs calibration technique is developed to measure the camera's pose and transform the 3C distances measured by each camera pair in their unique local camera coordinate system into the same local coordinate system. The camera pose is the camera's position and rotation at a given time point. This effort addresses *Challenge #1* of the dual stereo vision technique. For this process, only the left camera poses are used (the transformations from the right camera to the left camera are known through $\mathbf{A}_{r\to l}$ in Eq. (2), found during camera pair rectification). Since two camera pairs face different directions and cannot see the same ROI, a unique 3D calibration target is created. Specifically, a checkerboard calibration pattern with numbered squares is attached to two flat surfaces normal to each other. Two sheets of glass were attached to the wall and then ground and leveled to ensure they were perpendicular. A long checkerboard pattern was printed on thick paper and glued to the glass panes for a smooth surface. Next, the camera rig was suspended above the 3D calibration target, and a short, 30-second video was recorded while the camera rig was randomly rotated and translated by hand in all 6-DOFs. The set-up of this calibration process is shown in Figure 2.3. Since the distances between the corners of the checkerboard are known (i.e., each square has a width of 1.651 cm), the relationship between the two camera pairs was found. First, the transformation to convert the points measured in the local coordinate system of the camera to global coordinates [X Y Z 1]^T is found by

$$[i \quad j \quad k \quad 1]^T = [\mathbf{A}_{w\to c}^{ROI}][X \quad Y \quad Z \quad 1]^T \quad (3)$$

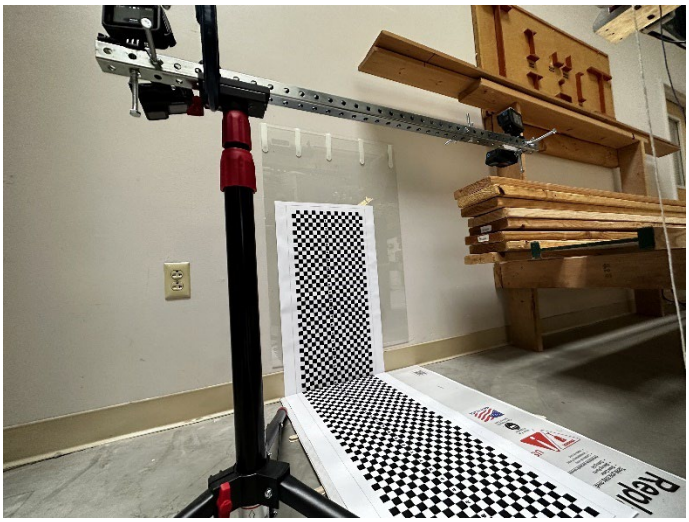


Figure 2.3 3D calibration target setup for dual stereo pairs calibration with the camera rig suspended above the target

$$[x \quad y \quad z \quad 1]^T = [\mathbf{A}_{w\to c}^{REF}][X \quad Y \quad Z \quad 1]^T \quad (4)$$

where $[i\ j\ k\ 1]^T$ are the identified points in the local camera coordinates of the *ROI* cameras, $[x\ y\ z\ 1]^T$ are the identified points in the local camera coordinates of the *REF* cameras, and $[\mathbf{A}^{ROI}{}_{w\to c}]$ and $[\mathbf{A}^{REF}{}_{w\to c}]$ are the transformation matrices from global coordinates to local camera coordinates for the *ROI* and *REF*

cameras, respectively. The transformation matrices comprise a 3 × 3 rotation matrix and 3 × 1 translation vector.

$$[\mathbf{A}_{w\to c}^{ROI}] = \begin{bmatrix} R_{w\to c}^{ROI} & T_{w\to c}^{ROI} \\ 0 & 1 \end{bmatrix}_{4\times 4} \quad (5)$$

$$[A_{w\to c}^{REF}] = \begin{bmatrix} R_{w\to c}^{REF} & T_{w\to c}^{REF} \\ 0 & 1 \end{bmatrix}_{4\times 4} \quad (6)$$

Since the coordinates of checkerboard points are known in both the local camera coordinates and the global coordinates, the transformation is solved using singular vector decomposition [36]. First, the covariance matrix, $\boldsymbol{H}$, is defined for each transformation matrix given a set of n points in a frame as

$$H^{ROI} = \left[\begin{bmatrix} i \\ j \\ k \end{bmatrix}_{3\times n} - \begin{bmatrix} \bar{\imath} \\ \bar{\jmath} \\ \bar{k} \end{bmatrix}_{3\times 1} \right] \times \left[\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{3\times n} - \begin{bmatrix} \bar{X} \\ \bar{Y} \\ \bar{Z} \end{bmatrix}_{3\times 1} \right]^T \quad (7)$$

$$H^{REF} = \left[\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{3\times n} - \begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix}_{3\times 1} \right] \times \left[\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{3\times n} - \begin{bmatrix} \bar{X} \\ \bar{Y} \\ \bar{Z} \end{bmatrix}_{3\times 1} \right]^T \quad (8)$$

where *n* is the number of identified points in one frame, and the barred coordinates represent the average coordinates of the *n* points. After the covariance matrices for the *ROI* and *REF* points are found, singular value decomposition of these covariance matrices is performed such that $\boldsymbol{H} = \boldsymbol{U} \times \boldsymbol{S} \times \boldsymbol{V}^T$. By multiplying the right singular vectors, $\boldsymbol{V}$, by the transpose of the left singular vectors, $\boldsymbol{U}$, the rotation matrix, $\boldsymbol{R_{w\to c}}$, is found.

$$[R_{w\to c}^{ROI}]_{3\times 3} = [V^{ROI}]_{3\times 3} \times [U^{ROI}]_{3\times 3}^T \quad (9)$$

$$[R_{w\to c}^{REF}]_{3\times 3} = [V^{REF}]_{3\times 3} \times [U^{REF}]_{3\times 3}^T \quad (10)$$

Lastly, the translation vectors are found by

$$[T_{w\to c}^{ROI}]_{3\times 1} = [\bar{X} \quad \bar{Y} \quad \bar{Z}]_{3\times 1}^T - [R_{w\to c}^{ROI}]_{3\times 3} \times [\bar{\imath} \quad \bar{\jmath} \quad \bar{k}]_{3\times 1}^T \quad (11)$$

$$[T_{w\to c}^{REF}]_{3\times 1} = [\bar{X} \quad \bar{Y} \quad \bar{Z}]_{3\times 1}^T - [R_{w\to c}^{REF}]_{3\times 3} \times [\bar{x} \quad \bar{y} \quad \bar{z}]_{3\times 1}^T \quad (12)$$

With the solved transformations, the transformation from the local coordinate system of the *ROI* camera to the local coordinate system of the *REF* camera, $[A^{ROI\to REF}]$, is found by converting the *ROI* points to global coordinates and then to the local *REF* coordinate system with the following transformation:

$$[A^{ROI\to REF}] = \begin{bmatrix} R_{w\to c}^{REF} & T_{w\to c}^{REF} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R_{w\to c}^{ROI} & T_{w\to c}^{ROI} \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} R^{ROI\to REF} & T^{ROI\to REF} \\ 0 & 1 \end{bmatrix}_{4\times 4} \quad (13)$$

Therefore, the points in the local camera coordinate system of the *ROI* camera are converted to the local camera coordinates of the *REF* camera by

$$[x \quad y \quad z \quad 1]^T = [A^{ROI\to REF}][i \quad j \quad k \quad 1]^T \quad (14)$$

With the above-defined transformation matrix, $[A^{ROI\to REF}]$, the points measured from the *ROI* camera pair (in engineering units in the *ROI* local camera coordinates) are then transformed into the local camera coordinates of the *REF* camera pair.

To reduce the error of this transformation matrix in Eq. (13), a short video is taken, and the transformation matrix of each frame is averaged over the video. Specifically, the rotation matrices of each frame, $[R^{ROI\rightarrow REF}]$ in Eq. (13), are averaged by converting the matrix into a quaternion, which is an alternate technique to describe rotations in 3D space using an ordered set of four numbers [37]. The quaternions are then averaged and converted back into the rotation matrix, $[R^{ROI\rightarrow REF}]$. The translation vector of each frame, $[T^{ROI\rightarrow REF}]$ in Eq. (13), is also averaged over the video. The averaged rotation matrix and translation vector are then used to define the transformation matrix, $[A^{ROI\rightarrow REF}]$. $[A_{w\rightarrow c}^{REF}]$ and $[A_{w\rightarrow c}^{ROI}]$ can vary from frame to frame depending on the movement of the cameras; however, $[A^{ROI\rightarrow REF}]$ remains constant since the position of the camera pairs relative to each other remains constant. With all the measured points in the same local coordinate system, the measured distances are related and later transformed into global coordinates for displacement measurements. This process is done only once for the two camera pairs and remains constant if the cameras are not shifted (i.e., during transportation of the camera rig).

2.3 Dual Stereovision Measurements

To provide high accuracy at a farther measuring distance, the baseline between the cameras within a camera pair is increased to about 80 cm. Theoretically, with an extended baseline, the accuracy of the measurements increases; however, there are practical factors that limit the baseline length in implementation, such as the physical constraint of the size of the camera rig that is feasible for a UAS to carry and the inability to match keypoints with a wide baseline due to different geometric and lightning conditions between images [38].

Each stereo vision camera pair measures the *x*-, *y*-, and *z*-coordinates of a matched keypoint in the local camera coordinate system of the left camera. The stereo vision-based measurement starts with matching keypoints, which are matched between images taken by the two cameras. To avoid using artificial patterns, natural features intrinsic in the material are found and tracked. Keypoints of natural features are initially manually identified within the image frame where measurements are needed. This minimal user input only requires the user to select matching keypoints in each camera pair for the first image frame. For instance, if the user wants to track a bolt's motion, he/she can manually select a keypoint on the bolt in the left camera and its corresponding keypoint in the right camera of the pair, forming a matching keypoint set. The procedure can be aided by an extremely close zooming in on the bolt to find a matching keypoint between the two cameras. An advantage of manually initializing and matching keypoints over other automatic techniques is that measurements can be taken at any location where a keypoint can be tracked instead of relying on the identified keypoint location with other algorithms. In this implementation, five keypoints were matched between the two cameras in each pair. Optical flow is then implemented to track the selected keypoints frame-to-frame. Optical flow tracks keypoints using a Harris Corner Detector and then searches within a small window in the next frame for a similar point [39], [40], [41]. Alternatively, the keypoints can be automatically matched by other techniques, such as block matching, which is typically used to generate depth maps of the entire FOV of the cameras in stereo vision applications [42]. However, manually matching the keypoints initially and implementing sub-pixel optical flow across the frames produced more precise results than using the block matching technique, making tracking natural features possible. Using the existing automated matching techniques, tracking the natural features can be unreliable. After the keypoints between the two cameras are found, the 3D locations of the keypoints are determined using stereo vision.

With a matched set of identified keypoints, the 3D locations of the keypoints are found. The variables for the stereo vision calculations are defined in Figure 2.4 and Figure 2.5. The image plane in Figure 2.4 and Figure 2.5 is the rectified image plane. Correspondingly, f (which is the same for both cameras) and B are the virtual focal length and baseline, respectively. Then, u_{pix}^{L} and v_{pix}^{L} are the pixel coordinates in

horizontal (left is the positive direction) and vertical directions (upward is the positive direction) with the virtual principal point of the left camera being the origin of the local pixel coordinate system, while u_{pix}^{R} is the pixel coordinate in horizontal direction (left is the positive direction) with the virtual principal point of the right camera being the origin of the pixel coordinate system. The star on the target is an identified and matched keypoint, and x, y, and z are the 3D coordinates in the local camera coordinate system. Using similar triangles, the 3D geometry is defined by,

$$\frac{f}{u_{pix}^{L}} = \frac{z}{x};\ \frac{f}{u_{pix}^{R}} = \frac{z}{(x-B)};\ \frac{f}{v_{pix}^{L}} = \frac{z}{y} \quad (15)$$

$$x = \frac{B*u_{pix}^{L}}{(u_{pix}^{R}-u_{pix}^{L})};\ z = \frac{B*f}{(u_{pix}^{R}-u_{pix}^{L})};\ y = \frac{B*v_{pix}^{L}}{\left(u_{pix}^{R}-u_{pix}^{L}\right)} \quad (16)$$

where $(u_{pix}^{L} - u_{pix}^{R})$ is defined as *disparity*, which is the difference of the pixel coordinate of a pixel in the left camera to the pixel coordinate of the corresponding pixel in the right camera.

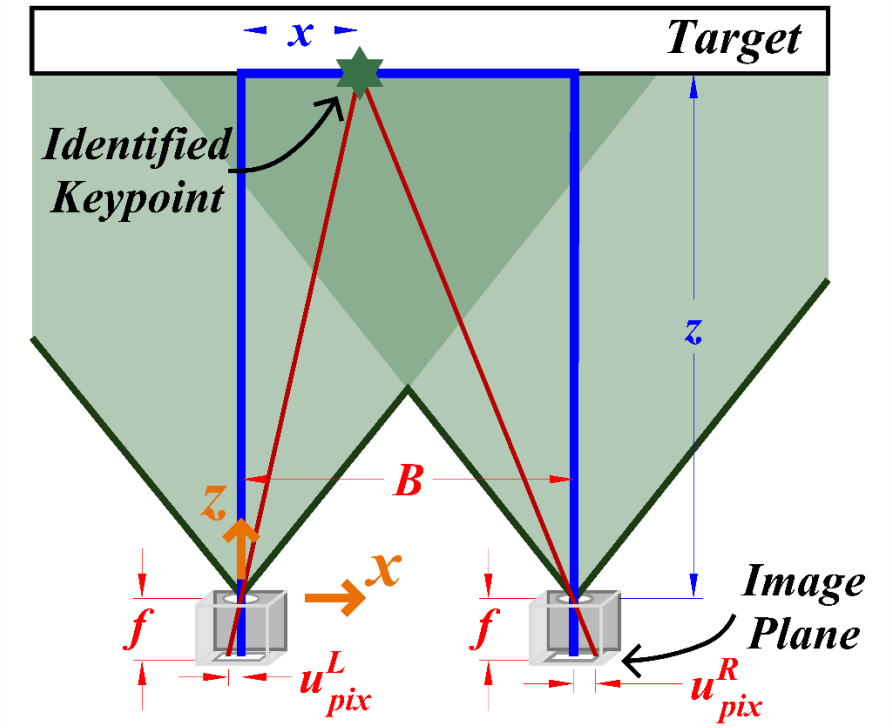


Figure 2.4 Schematic diagram of the general stereovision geometry (plotted based on REF cameras): solving for the x- and z- directions

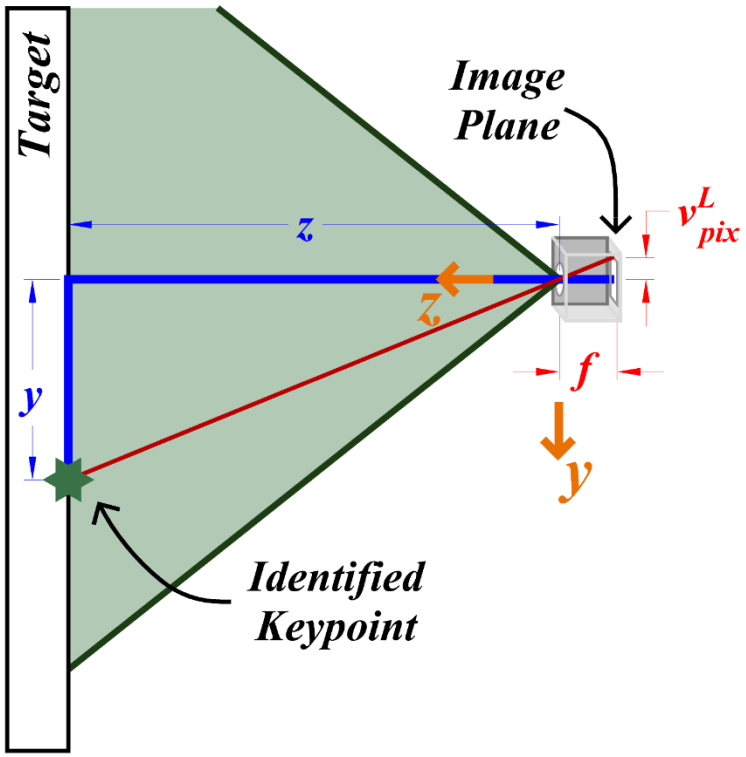


Figure 2.5 Schematic diagram of the general stereovision geometry (plotted based on REF cameras): solving for the *y*-direction

Lastly, to address *Challenge #2*, the translation and rotation of the UAS are measured on a frame-by-frame basis. The rotation and translation measurement of the UAS follows a similar mathematical principle to that stated in Sabharwal and Guo [43] for tracking windborne debris in space using stereo vision. The technique proposed in this study directly solves for the transformation matrix from the global coordinate system to the local coordinate system of the left REF camera. Using this transformation matrix, the 6-DOF motion of the UAS is recovered by tracking points in a reference plane. The directional unit vectors, $\hat{\boldsymbol{u}}$, $\hat{\boldsymbol{v}}$, and $\hat{\boldsymbol{w}}$, and the center point, $\boldsymbol{C}$, are defined using the reference plane in Figure 2.6. First, a plane is fitted to points $\boldsymbol{Q1}$ through $\boldsymbol{Q5}$ in Figure 2.6 using least squares. This fitted plane is the reference plane in the real world. The reference points are split into two groups: group 1 consists of $\boldsymbol{Q1}$, $\boldsymbol{Q2}$, and $\boldsymbol{Q3}$, and group 2 consists of $\boldsymbol{Q3}$, $\boldsymbol{Q4}$, and $\boldsymbol{Q5}$. The centroids of the two groups are found and projected onto the fitted reference plane (i.e., $\boldsymbol{C}$ and $\boldsymbol{D}$ for group 1 and group 2, respectively). The directional vector, $\hat{\boldsymbol{w}}$, is found by taking the unit normal of the fitted plane. Next, the directional vector, $\hat{\boldsymbol{u}}$, is defined as the unit vector from $\boldsymbol{C}$ to $\boldsymbol{D}$: $\hat{\boldsymbol{u}} = \frac{\boldsymbol{CD}}{|\boldsymbol{CD}|}$. Lastly, the directional vector, $\hat{\boldsymbol{v}}$, is defined as the cross product of $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{w}}$: $\hat{\boldsymbol{v}} = \frac{\hat{\boldsymbol{u}}\times\hat{\boldsymbol{w}}}{|\hat{\boldsymbol{u}}\times\hat{\boldsymbol{w}}|}$. Point $\boldsymbol{C}$ is defined as the origin of the global coordinate system, while $\hat{\boldsymbol{u}}$, $\hat{\boldsymbol{v}}$, and $\hat{\boldsymbol{w}}$ define the mutually perpendicular coordinate axis. Note that only three non-collinear points in the reference plane are required; however, it was found that using two groups of three points produced a more stable global coordinate system as it averaged out random fluctuations of the point measurements. The transformation matrix from global points to the local camera coordinates of the REF camera is defined by,

$$[\boldsymbol{F}] = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{C} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{u}} & \hat{\boldsymbol{v}} & \hat{\boldsymbol{w}} & \boldsymbol{C} \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \hat{u}_1 & \hat{v}_1 & \hat{w}_1 & C_1 \\ \hat{u}_2 & \hat{v}_2 & \hat{w}_2 & C_2 \\ \hat{u}_3 & \hat{v}_3 & \hat{w}_3 & C_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (17)$$

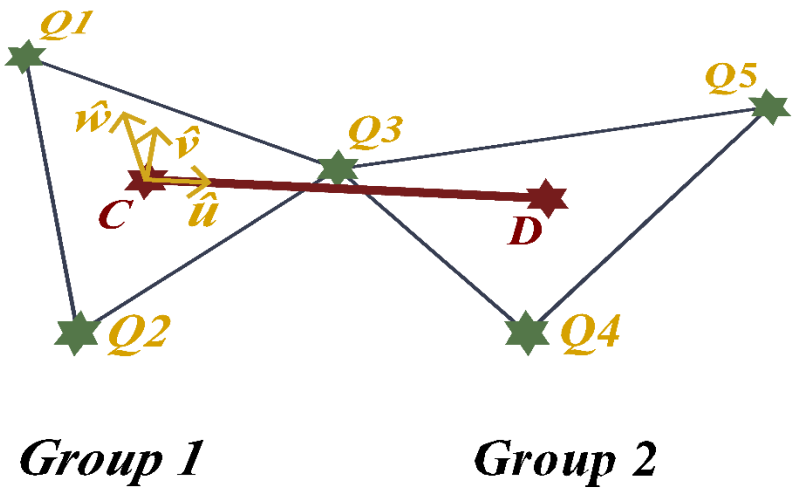


Figure 2.6 Reference plane and the global coordinate system

The inverse of $[\boldsymbol{F}]$, $[\boldsymbol{F}^{-1}]$, defines the transformation from the local camera coordinates of the REF camera to the global coordinate system,

$$[\boldsymbol{F}^{-1}] = \begin{bmatrix} \hat{\boldsymbol{u}}^T & -\hat{\boldsymbol{u}}^T\boldsymbol{C} \\ \hat{\boldsymbol{v}}^T & -\hat{\boldsymbol{v}}^T\boldsymbol{C} \\ \hat{\boldsymbol{w}}^T & -\hat{\boldsymbol{w}}^T\boldsymbol{C} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \hat{u}_1 & \hat{v}_1 & \hat{w}_1 & -(\hat{u}_1C_1 + \hat{u}_2C_2 + \hat{u}_3C_3) \\ \hat{u}_2 & \hat{v}_2 & \hat{w}_2 & -(\hat{v}_1C_1 + \hat{v}_2C_2 + \hat{v}_3C_3) \\ \hat{u}_3 & \hat{v}_3 & \hat{w}_3 & -(\hat{w}_1C_1 + \hat{w}_2C_2 + \hat{w}_3C_3) \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (18)$$

$[\boldsymbol{F}^{-1}]$ allows the transformation of points in the local coordinate system of the REF camera into global coordinates. Since the reference points do not move, the reference coordinate system remains constant. This transformation is solved on a frame-by-frame basis to compensate for the random drift and rotation of the UAS.

With all the transformation matrices defined, the points measured in the local camera coordinate system of the ROI cameras ($[i \quad j \quad k \quad 1]^T$) are directly converted into the constant global coordinates ($[X \quad Y \quad Z \quad 1]^T$), where the 6-DOF of UAS motion is compensated for via $[\boldsymbol{F}^{-1}]$. Specifically, the ROI points in the local ROI camera coordinate system are first converted into the local REF camera coordinate system through $[\boldsymbol{A}^{ROI\rightarrow REF}]$. The points in the local REF camera coordinate system are then transformed into global coordinates through $[\boldsymbol{F}^{-1}]$. Finally, the measured 3C motion of ROI points is represented in the constant global coordinates by

$$[X \quad Y \quad Z \quad 1]^T_{4\times n} = [F^{-1}]_{4\times4}[A^{ROI\rightarrow REF}]_{4\times4}[i \quad j \quad k \quad 1]^T_{4\times n} \quad (19)$$

3. EXPERIMENTAL RESULTS

3.1 Validation of the Proposed Dual Stereovision Technique with Nonstationary Cameras

The results of the proposed technique validation are presented in this section. The full calibration process was performed once and assumed to be constant throughout the rest of the experiments. Each camera pair was calibrated and rectified using MATLAB's built-in Stereo Camera Calibrator App. The pixel error of this calibration process was less than 0.5 pixels for both camera pairs. After calibration and rectification, the ROI and REF camera pairs underwent the dual stereo pairs calibration procedure to find the transformation between the coordinate systems of the two camera pairs. Then, the transformation from the local coordinate system of the REF camera to the global coordinate system is solved for each frame. To test the efficacy of the entire workflow, an experiment is conducted to measure the global position of stationary points on the 3D calibration target (Figure 2.3) while the camera rig is moving. This experiment aims to test if the proposed camera system can accurately measure the 3D locations of the tracked stationary points in the global coordinate system when the camera rig constantly moves with a significant amount of random rotation and translation during the video recording. This experiment serves as a first step in validating the proposed algorithm for simultaneously measuring the locations of real-world targets and 6-DOF camera rig motion. Specifically, in this experiment, the camera rig was significantly and randomly rotated and translated by hand above the 3D calibration planes while the cameras were tracking the relative motion between the camera and the stationary points via videos. Although the cameras were moving during the test, after applying the proposed transformation matrices, the resolved locations of tracked points should stay constant over time as those points were stationary in the global coordinate system. However, there is a slight error likely attributed to tracking errors, imperfect calibrations, and imperfect synchronization of the cameras.

To synchronize the cameras, initially, remote triggering of the cameras using GoPro's proprietary remote was used; however, it was observed from many repetitive experiments that if the cameras were out of sync by more than half of a frame (about 0.010 seconds), the resulting measurement error became unacceptable (the frequency information of the response cannot be recovered). Therefore, the camera frame offsets were measured through post-processing. At the beginning of the video, a few sharp sounds (i.e., hand claps) were generated, and the autocorrelation functions of the sound waves between the cameras were measured. Since the sound waves have a sampling rate of 48 kHz, which is much higher than the sampling rate of image frames (60 Hz), the precise frame time offset was found. Next, the measured keypoint positions of each camera were linearly interpolated between the frames to achieve sub-frame time synchronization.

The plot of the estimated X, Y, and Z location of a stationary point is shown for 1,200 frames (about 20 seconds) in Figure 3.1, Figure 3.2, and Figure 3.3. The translation and rotation of the camera rig during this test are also plotted in Figure 3.4 and Figure 3.5. Note the significant translation and rotation of the camera rig (on the order of ± 50 mm and $\pm 15°$) during this test. The root-mean-square (RMS) error to the ground truth is 0.61 mm, 0.89 mm, and 0.63 mm for the X-, Y-, and Z-directions, respectively. The sub-millimeter RMS error shows that the mathematical concept of dual camera pairs calibration, stereo vision measurement, and compensation of 6-DOF motion of moving cameras is valid in practice.

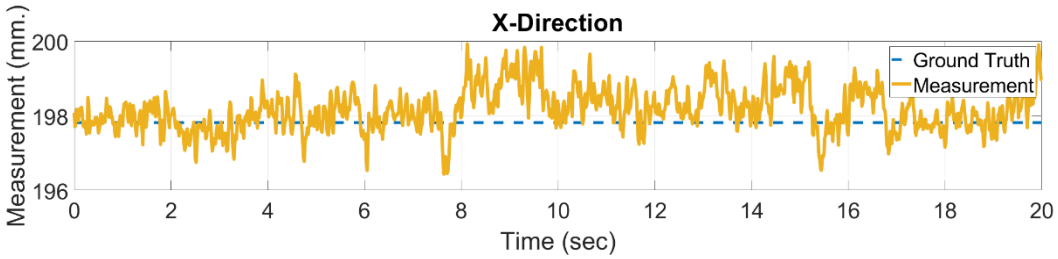


Figure 3.1 Measured location of a stationary point using moving dual stereo camera pairs in the global *x*-coordinate

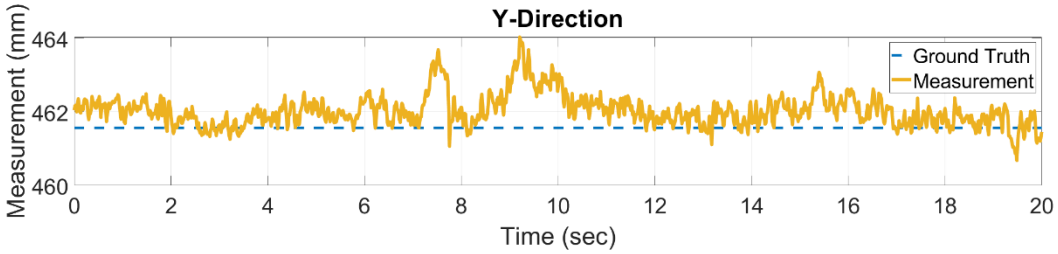


Figure 3.2 Measured location of a stationary point using moving dual stereo camera pairs in the global *y*-coordinate

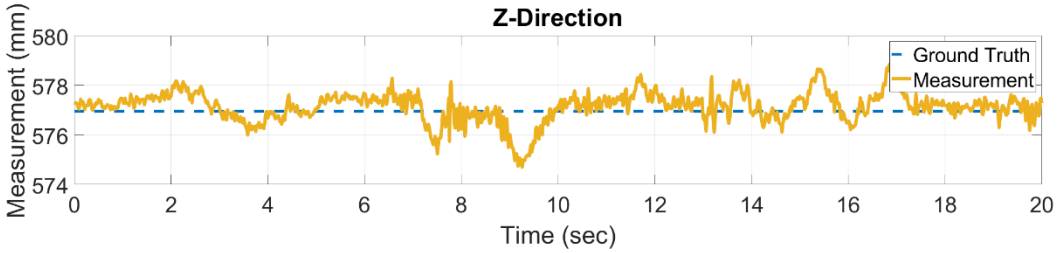


Figure 3.3 Measured location of a stationary point using moving dual stereo camera pairs in the global *z*-coordinate

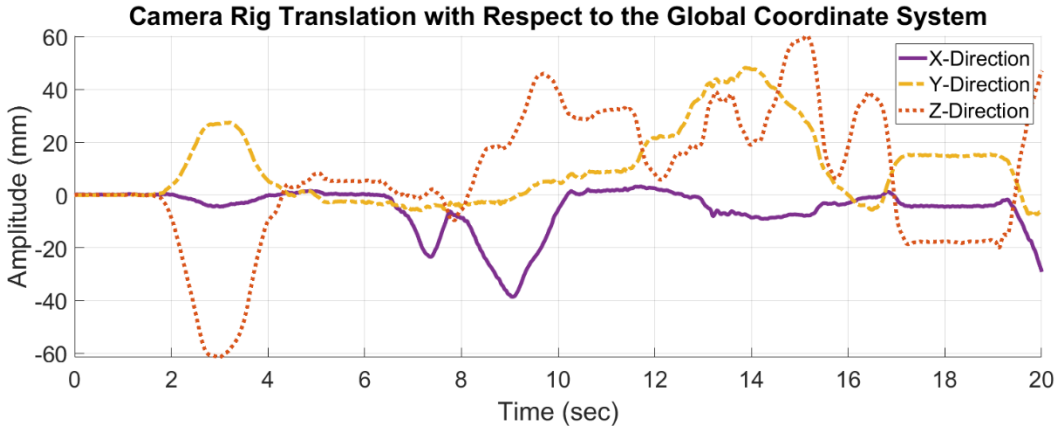


Figure 3.4 The captured motion of the camera rig during the stationary point test: 3-DOF translation of the camera rig

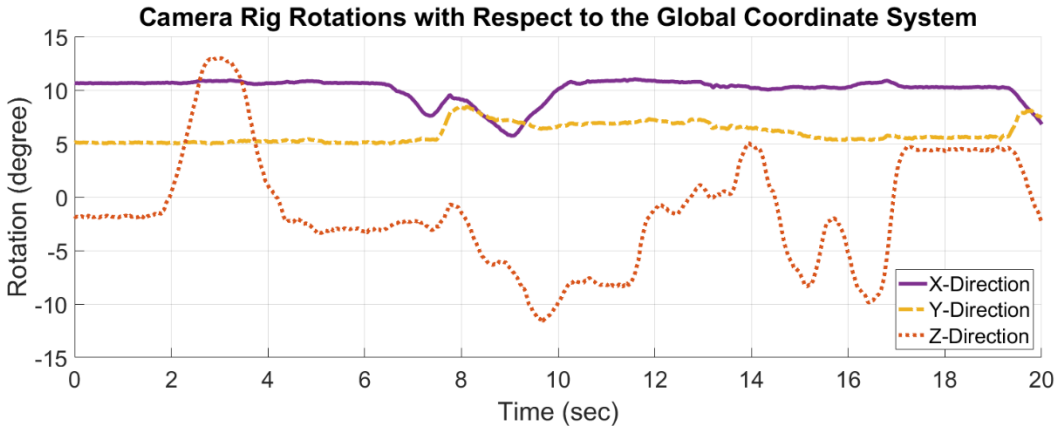


Figure 3.5 The captured motion of the camera rig during the stationary point test: 3-DOF rotation of the camera rig

3.2 UAS-Based Measurement Results

After the proposed technique was validated in a lab setting with high accuracy, the camera rig was placed on a UAS and flown to test the efficacy of measuring dynamic displacements. A shake table that simulated sine waves in two directions (*X*- and *Y*-directions) was used to provide a dynamic ROI with known motion. Two concrete beams were placed on top of the shake table to simulate natural features that would be present in a concrete structure. The frame rate of the cameras is 60 fps. The UAS was flown approximately 5 m from the shake table. This distance is believed to be sufficient for field applications. The experimental setup is shown in Figure 3.6. Checkered patterns were placed on the ground under the UAS and on the concrete on the shake table to provide a redundant data source to help debug and verify the algorithm. However, these patterns were not used in the final data processing [Eqs. (10) and (11)], and the presented results are only from using natural features in the concrete and on the ground. The tracked points from natural features used in the final test are shown in Figure 3.7 and Figure 3.8. These natural points on the concrete specimen and the ground were selected manually using the methodology described in the Dual Stereo Vision Measurements section.

The shake table's motion directions did not perfectly align with the reference coordinate system on the ground (Figure 3.6). To be able to compare the ground truth from the shake table with the motion measured with respect to the reference coordinate system, the reference coordinate system was transformed to align with the motion of the shake table. The transformation matrix was solved by finding the dominant direction of motion of the measured data through singular value decomposition such that $\boldsymbol{G} = \boldsymbol{U} \times \boldsymbol{S} \times \boldsymbol{V}^T$, with

$$\boldsymbol{G} = \left[\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{3\times n} - \begin{bmatrix} \bar{X} \\ \bar{Y} \\ \bar{Z} \end{bmatrix}_{3\times 1} \right] \times \left[\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{3\times n} - \begin{bmatrix} \bar{X} \\ \bar{Y} \\ \bar{Z} \end{bmatrix}_{3\times 1} \right]^T \quad (20)$$

The dominant direction is described by the right singular vector, $\boldsymbol{V_1}$, corresponding to the largest singular value, S_1. Knowing the dominant direction, the transformation matrix is solved directly using Eq. (17). Specifically, the dominant direction forms $\hat{\boldsymbol{u}}$ in Eq. (17), while $\hat{\boldsymbol{v}}$ and $\hat{\boldsymbol{w}}$ are determined using the same approach described in the Dual Stereo Vision Measurements section. Lastly, a third-order high-pass Butterworth filter with a stop band of 0.1 Hz and a third-order low-pass Butterworth filter with a stop band of 10 Hz were applied to the measurements. Both filters were implemented using the "butter" function in Matlab. The bandwidth of the filter does not affect the shake table signal. These two filters

removed the high-frequency variations due to the slightly inaccurate subpixel tracking and the low-frequency variations due to the non-compensated UAS motion. The results for the test are shown in Figure 3.9, Figure 3.10, and Figure 3.11 with the ground truth signal and tabulated in Table 3.1. Note that the shake table did not move in the *Z*-direction. The RMS errors of the measured motion and the ground truth are 2.54 mm, 5.77 mm, and 6.40 mm in the global coordinate system in the *X*-, *Y*-, and *Z*-directions, respectively. In addition, the time histories of 3-DOF translational motion and 3-DOF rotation of the UAS with respect to the global coordinate system defined in Figure 3.6 are plotted in Figure 3.12, Figure 3.13, Figure 3.14, and Figure 3.15. Note that the measured dynamic displacement (Figure 3.9, Figure 3.10, and Figure 3.11) and the captured motion of the UAS (Figure 3.12, Figure 3.13, Figure 3.14, and Figure 3.15) show little correlation, demonstrating the motions of the UAS are removed successfully from the measured displacements. Also note that the magnitude of the UAS motion, which is estimated by taking the difference between the maximum and minimum values of each time history of displacement in Figure 3.12, Figure 3.13, Figure 3.14, and Figure 3.15, is quite significant. Specifically, the magnitudes of UAS translation in the *X*-, *Y*-, and *Z*-direction are on the order of 200 mm, 250 mm, and 500 mm, respectively. The magnitudes of UAS rotation with respect to the *X*-, *Y*- and *Z*-axis are 2°, 5°, and 2°, respectively. Note that the concrete beam's magnitude of motion is only 10 to 50 mm. Measuring the relatively small amplitude of displacements with the significant UAS motion has demonstrated the prowess of the proposed technique. The power spectral densities of the measured 3C displacements are shown in Figure 3.16.

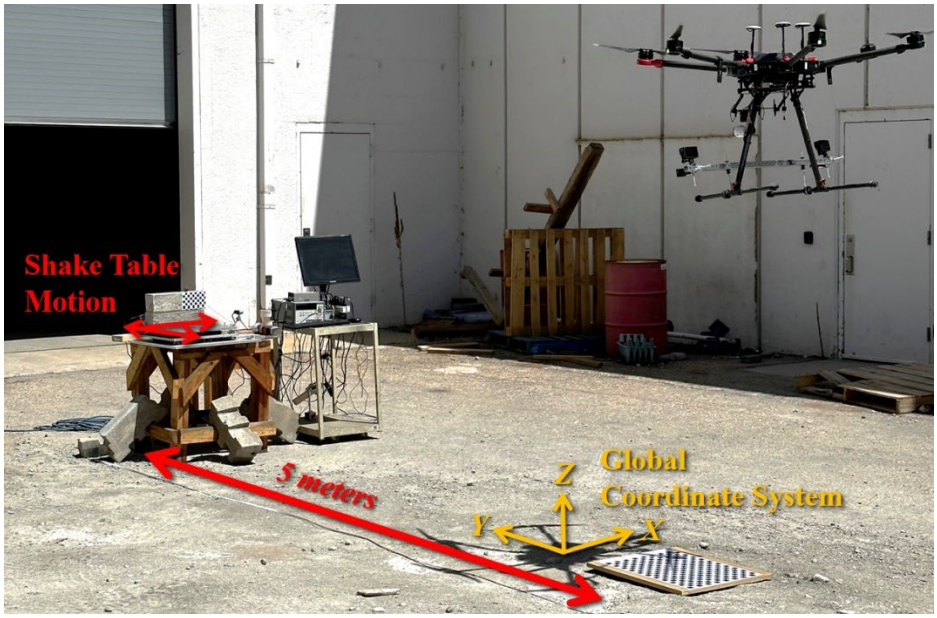


Figure 3.6 UAS-based lab experiment with shake table motion

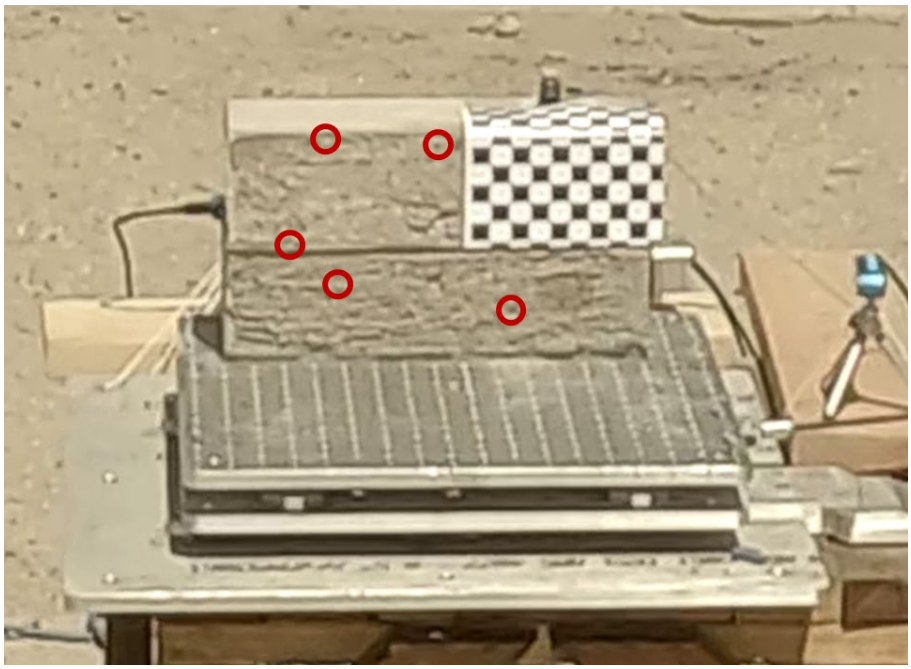


Figure 3.7 Tracked natural features for the ROI measurement area

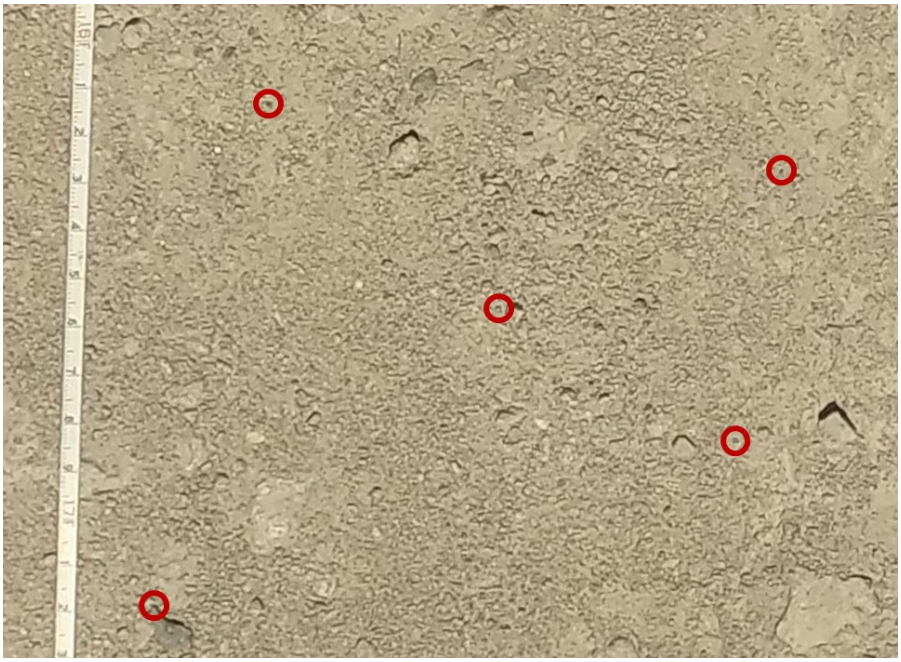


Figure 3.8 Tracked natural features for the REF measurement area

The percent differences (absolute values) of the estimated frequencies to the ground truth are shown in Table 3.1. In Table 3.1, the frequency percent difference in *Z*-direction does not apply as there was no motion of the specimen in *Z*-direction. The dominant frequency content around 0.6 Hz seen in *Z*-direction in Figure 3.16 is spurious due to measurement noise. Note that due to the relatively short signals (less than 10 seconds), the errors in frequency estimation of *X*-, and *Y*-direction are partially attributed to the short window effect on the Fourier transform. From Figure 3.9, Figure 3.10, and Figure 3.11, there is slightly more measurement noise (manifested as jagged waveforms) in the *X*- and *Z*-directions than the *Y*-

direction. This is attributed to the nature of the stereo-vision technique, in which both the x- and y-direction measurements are dependent on the z-direction measurement and thus may contain additional errors propagated from z-direction measurement [Eq. (15)]. Note that the global coordinate system in Figure 3.9, Figure 3.10, and Figure 3.11 roughly aligns with the local coordinate system in such a way that $x \sim X$, $y \sim Z$, and $z \sim Y$. Thus, it is expected to see slightly more errors in X- and Z-direction measurements in Figure 3.9, Figure 3.10, and Figure 3.11 compared with the Y-direction measurement.

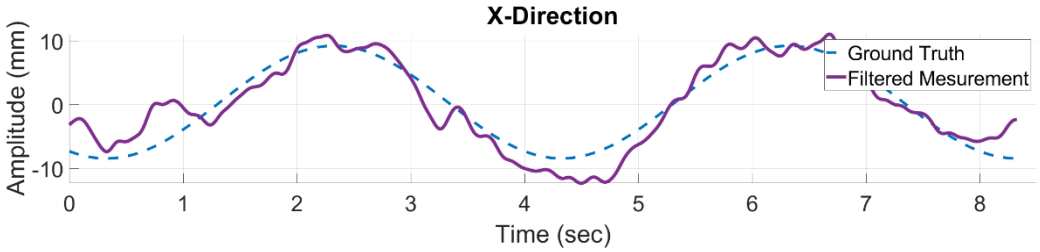


Figure 3.9 Comparison of the measured motion of the concrete on the shake table against the ground truth measurements in the x-direction

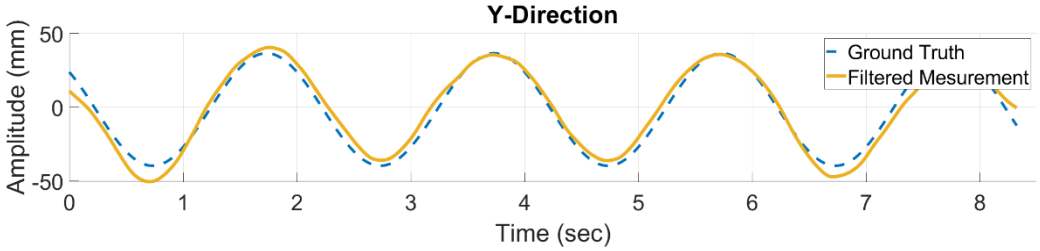


Figure 3.10 Comparison of the measured motion of the concrete on the shake table against the ground truth measurements in the y-direction

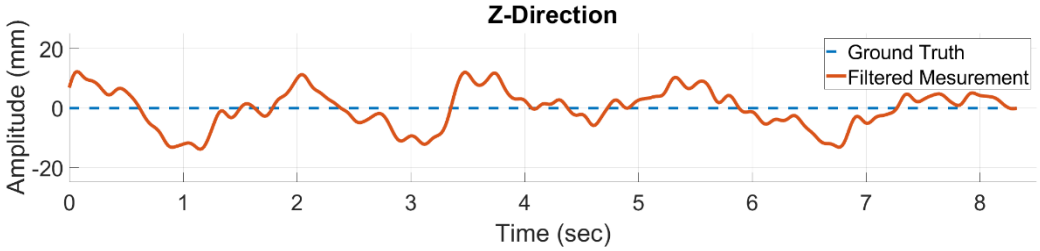


Figure 3.11 Comparison of the measured motion of the concrete on the shake table against the ground truth measurements in the z-direction

Table 3.1 Results of UAS-based 3C dynamic displacements measurements

Displacement Component	UAS-based measurement RMS Error (mm)	UAS-based frequency difference (%)
X-direction	2.54	5.86
Y-direction	5.77	5.67
Z-direction	6.40	---

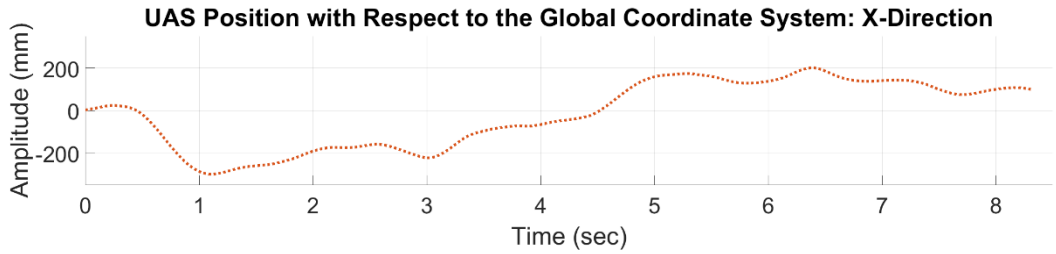


Figure 3.12 Captured motion of the UAS with respect to the global coordinate system during flight: 3-DOF position in the *x*-direction

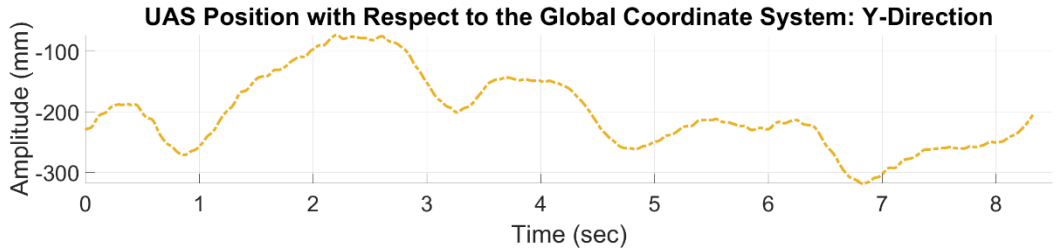


Figure 3.13 Captured motion of the UAS with respect to the global coordinate system during flight: 3-DOF position in the *y*-direction

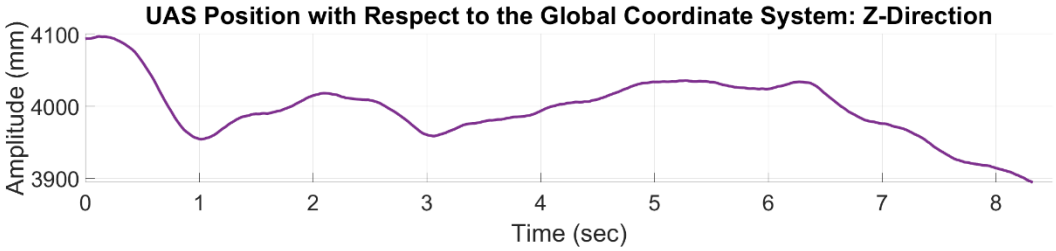


Figure 3.14 Captured motion of the UAS with respect to the global coordinate system during flight: 3-DOF position in the *z*- direction

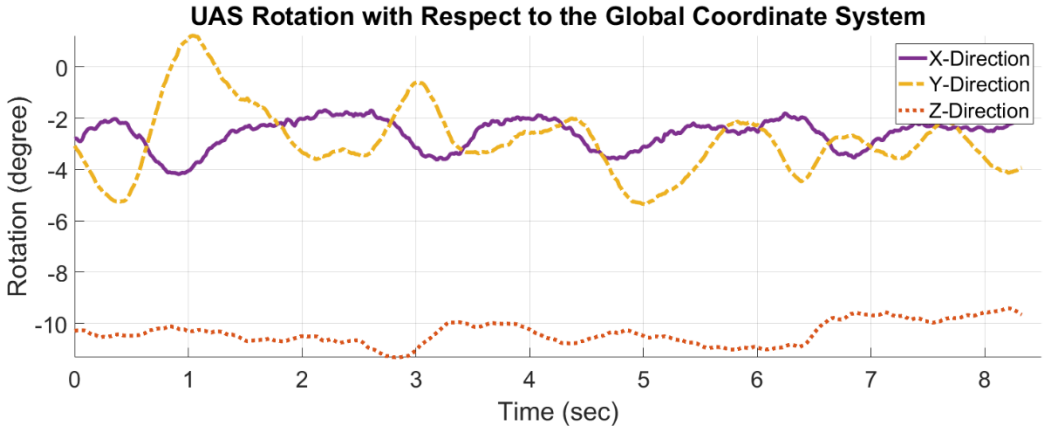


Figure 3.15 Captured motion of the UAS with respect to the global coordinate system during flight: 3-DOF rotations

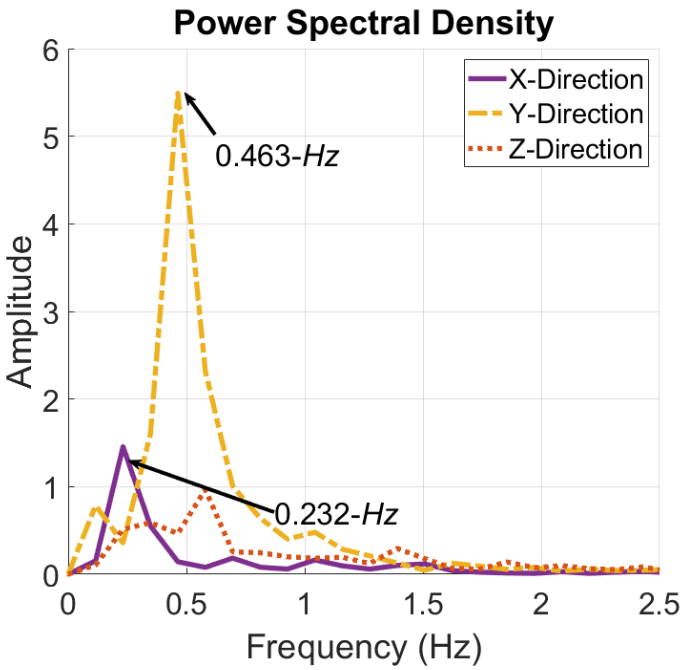


Figure 3.16 Power spectral density of the UAS-enabled 3C dynamic displacement measurements

4. DISCUSSIONS

Measurements of 3C dynamic displacements were taken of a concrete ROI using cameras attached to a UAS. The proposed system demonstrated sub-millimeter accuracy when measuring the positions of a stationary point, but errors were increased when the camera rig was flown farther from the ROI in the dynamic field test. Moving from the stationary test to the dynamic field test, there were three significant changes: 1) the use of natural features intrinsic to the concrete, 2) the dynamic motion of the ROI, and 3) a farther distance between the ROI and UAS. The potential error sources that may stem from each of these changes are analyzed. By comparing the results of the dynamic field test when tracking natural features and artificial patterns (i.e., the checkerboard pattern), the robustness of optical flow for tracking natural features in concrete was verified. Thus, there were no significant error sources due to tracking natural features. Unless there is motion blur or spatial distortion caused by high relative speeds between the ROI and UAS, the motion of the ROI is not expected to cause measurement errors either. Since the frame rate was high (60 Hz) and the relative speed between the ROI and UAS was low, no motion blur or spatial distortions were observed in the video, and optical flow performed as expected. Therefore, the dynamic motion of the ROI is not expected to cause measurement errors either. 3) In comparing the dynamic test with the stationary test, there was an almost 10-fold increase in the GSD (from 0.20 mm/pix to 1.99 mm/pix). There was also about a 10-fold increase in the measurement error (from 0.63 mm to 6.40 mm in the *Z*-direction). These results suggest that the increased GSD may have caused the increased measurement error. Note that the higher value of GSD indicates lower spatial resolution of the image and fewer visible details. An increased GSD at a farther camera-object distance may lead to larger measurement errors due to reduced visibility of details in the object. Thus, the increased error observed in the dynamic test is likely partially attributed to the increased GSD at the greater working distance. To improve the accuracy of measurements at farther distances, one can implement methods to reduce the GSD. The GSD is a function of the number of pixels the camera can capture, the sensor size, and the focal length. Reducing the GSD can be achieved physically by using cameras with a larger number of pixels or lenses with longer focal lengths. In addition, artificial means to increase the camera resolution, such as super-resolution [44], could potentially reduce the GSD. These artificial means could be investigated in future research.

Based on the lessons learned from this study, there are some considerations for future improvements and applications:

(1) The pixel errors from camera calibration, keypoint tracking, and inaccurate camera syncing are the major sources of error that propagate to measurement errors in engineering units. Therefore, the performance of all three processes directly affects the accuracy of displacement measurements.
(2) It is important to ensure that the relative locations of the four cameras do not change after camera calibration, as the change of relative locations can invalidate the camera calibration and cause measurement errors. This requirement can be achieved through a properly designed rigid camera mounting system. The prototype camera rig used in the current study has strongly informed the feasibility of such a camera mounting system. Future studies will focus on designing a more permanent and durable camera mounting system to facilitate more general field applications.
(3) Synchronization of the cameras is an important consideration. Based on repeated experiments, the errors become unacceptable (the frequency information of the response cannot be recovered) if the cameras are out of sync by half of a frame (about 0.010 seconds in this study). Whenever possible, it is recommended to implement hardware synchronization of the cameras.
(4) It is important to estimate the GSD of the camera systems before any application, as the GSD is closely related to the expected error.

5. CONCLUDING REMARKS

This study presents a novel dual stereo vision technique to directly measure the dynamic displacements of a structure with a UAS platform. With two pairs of cameras attached to a UAS, the 3C dynamic displacement of the structure and the 6-DOF motion of the UAS are simultaneously measured using a mathematically elegant workflow. The following conclusions are drawn from this study.

(1) A novel pair calibration technique is developed to precisely determine the relative locations of the four cameras for the challenging condition when the same features are not visible in all four cameras.
(2) Using two pairs of cameras allows flexibility in choosing a stationary reference. The two camera pairs can be oriented in different directions to provide the best vantage points for both the ROI and reference, greatly enhancing the feasibility of applying the new technology in various field environments.
(3) The 3-DOF translation and 3-DOF rotation motion of the camera rig is successfully recovered by finding a transformation matrix from the global coordinate system to the local coordinate system of REF cameras.
(4) Natural features can be identified using minimal user input and tracked, providing the ability to take measurements anywhere without the use of artificial targets, which makes the proposed technique viable for difficult-to-access structures.
(5) The efficacy of the proposed technique was demonstrated via experiments, where the RMS errors of the measured displacement time histories at a 5-m distance from the cameras to the structure are 2.54 mm and 5.77 mm in the X- and Y-directions of the global coordinate system. This accuracy level was achieved when measuring displacements with amplitude in the order of 10 to 50 mm under UAS motion in the order of $\pm$250 mm (translation) and $\pm$5° (rotation). The ability of this technique to measure small-amplitude dynamic displacement with significant UAS motion allows this UAS-based remote sensing technique to be applied to civil structures under challenging environmental conditions such as high winds.
(6) The 5-m camera-to-structure distance achieved through this study is considered a significant improvement compared with that of the existing study (0.75 m) [30]. The area of the ROI captured by the cameras at this 5-m distance is 5.9 m $\times$ 10.6 m. This working distance and camera field of view have demonstrated the feasibility and applicability of the proposed technique on a large scale for many civil structures.

REFERENCES

[1] Li, Z., and C. Chang. 2012. "Tracking of structural dynamic characteristics using recursive stochastic subspace identification and instrumental variable technique." *Journal of Engineering Mechanics.* 138 (6): 591-600. https://doi.org/10.1061/(ASCE)EM.1943-7889.0000370.

[2] Rainieri, C., and G. Fabbrocino. 2010. "Automated output-only dynamic identification of civil engineering structures." *Mechanical Systems and Signal Processing.* 24 (3): 678-695. https://doi.org/10.1016/j.ymssp.2009.10.003.

[3] Brincker, R., L. Zhang, and P. Andersen. 2000. "Modal identification from ambient responses using frequency domain decomposition." In Proceedings of the International Modal Analysis Conference. San Antonio, TX: Aalborg Universitet.

[4] Guo, Y., A. Kareem, Y. Ni, and W. Liao. 2012. "Performance evaluation of Canton Tower under winds based on full-scale data." *Journal of Wind Engineering and Industrial Aerodynamics.* 104-106(Jun): 116-128. https://doi.org/10.1016/j.jweia.2012.04.001.

[5] Guo, Y., D. K. Kwon, and A. Kareem. 2016. "Near-real-time hybrid system identification framework for civil structures with application to Burj Khalifa." *Journal of Structural Engineering.* 142 (2): 04015132. https://doi.org/10.1061/(ASCE)ST.1943-541X.0001402.

[6] Guo, Y., and A. Kareem. 2016. "Non-stationary frequency domain system identification using time-frequency representations." *Mechanical Systems and Signal Processing.* 72-73 (5): 712-726. https://doi.org/10.1016/j.ymssp.2015.10.031.

[7] Guo, Y., and A. Kareem. 2015. "System identification through nonstationary response: Wavelet and transformed singular value decomposition-based approach." *Journal of Engineering Mechanics.* 141 (7): 04015013. https://doi.org/10.1061/(ASCE)EM.1943-7889.0000905.

[8] Kromanis, R., and P. Kripakaran. 2017. "Data-driven approaches for measurement interpretation: Analysing integrated thermal and vehicular response in bridge structural health monitoring." *Advanced Engineering Informatics.* 34 (10): 46-59. https://doi.org/10.1016/j.aei.2017.09.002.

[9] Moschas, F., and S. Stiros. 2011. "Measurement of the dynamic displacements and of the modal frequencies of a short-span pedestrian bridge using GPS and an accelerometer." *Engineering Structures.* 33 (1): 10-17. https://doi.org/10.1016/j.engstruct.2010.09.013.

[10] Roberts, G., X. Meng, and A. Dodson. 2004. "Integrating a Global Positioning System and Accelerometers to Monitor the Deflection of Bridges." *Journal of Surveying Engineering.* 130 (2): 65-72. https://doi.org/10.1061/(ASCE)0733-9453(2004)130:2(65).

[11] Kijewski-Correa, T., A. Kareem, and M. Kochly. 2006. "Experimental Verification and Full-Scale Deployment of Global Positioning Systems to Monitor the Dynamic Response of Tall Buildings." *Journal of Structural Engineering.* 132 (8): 1242-1253. https://doi.org/10.1061/(ASCE)0733-9445(2006)132:8(1242).

[12] Brown, R., K. Gaebler, C. Shield, and L. Linderman. 2019. "Displacement Monitoring of I-35W Saint Anthony Falls Bridge with Current Vibration-Based System." Rep. No. MN/RC 2019-05. Minneapolis, MN: Minnesota Department of Transportation.

[13] Kijewski-Correa, T., D. Kwon, A. Kareem, A. Bentz, Y. Guo, S. Bobby, and A. Abdelrazaq. 2013. "Smartsync: An integrated real-time structural health monitoring and structural identification system for tall buildings." *Journal of Structural Engineering.* 139 (10): 1675-1687. https://doi.org/10.1061/(ASCE)ST.1943-541X.0000560.

[14] Havaran, A., and M. Mahmoudi. 2020. "Extracting structural dynamic properties utilizing close photogrammetry method." *Measurement* 150 (1): 107092. https://doi.org/10.1016/j.measurement.2019.107092.

[15] Liu, X., X. Tong, W. Lu, S. Liu, B. Huang, P. Tang, and T. Guo. 2020. "High-speed videogrammetric measurement of the deformation of shaking table multi-layer structures." *Measurement* 154 (15): 107486. https://doi.org/10.1016/j.measurement.2020.107486.

[16] Dong, C., X. Ye, and T. Jin. 2018. "Identification of structural dynamic characteristics based on machine vision technology." *Measurement* 126 (10): 405-416. https://doi.org/10.1016/j.measurement.2017.09.043.

[17] Aghlara, R., and M. M. Tahir. 2018. "Measurement of strain on concrete using an ordinary digital camera." *Measurement* 126 (10): 398-404. https://doi.org/10.1016/j.measurement.2018.05.066.

[18] Kromanis, R., and C. Forbes. 2019. "A low-cost robotic camera system for accurate collection of structural response." *Inventions* 4 (3): 47. https://doi.org/10.3390/inventions4030047.

[19] Mousa, M., M. Yussof, U. Udi, F. Nazri, M. Kamarudin, G. Parke, L. Assi, and S. Ghahari. 2021. "Application of digital image correlation in structural health monitoring of bridge infrastructures: A review." *Infrastructures* 6 (12): 176. https://doi.org/10.3390/s23010470.

[20] Kromanis, R., and P. Kripakaran. 2021. "A multiple camera position approach for accurate displacement measurement using computer vision." *Journal of Civil Structural Health Monitoring.* 11 (3): 661-678. https://doi.org/10.1007/s13349-021-00473-0.

[21] Dong, C., S. Bas, and F.N. Catbas. 2023. "Applications of computer vision-based structural monitoring on long-span bridges in Turkey." *Sensors* 23 (19): 8161. https://doi.org/10.3390/s23198161.

[22] Celik, O., C.-Z. Dong, and F.N. Catbas. 2020. "Computer vision-based human comfort assessment of stadiums." *Journal of Performance of Constructed Facilities.* 34 (2): 04020005. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001345.

[23] Liu, X., S. J. Dyke, A. Lenjani, I. Bilionis, X. Zhang, and J. Choi. 2023. "Automated image localization to support rapid building reconnaissance in large-scale area." *Computer-Aided Civil and Infrastructure Engineering.* 38 (1): 3-25. https://doi.org/10.1111/mice.12828.

[24] Catt, S., B. Fick, M. Hoskins, J. Praski, and J. Baquersad. 2019. "Development of a semi-autonomous drone for structural health monitoring of structures using digital image correlation (DIC)." In Proceedings of the 36th IMAC, A Conference and Exposition on Structural Dynamics 2018,

[25] Yoon, H., V. Hoskere, J. Park, and B. Spencer. 2017. "Cross-correlation-based structural system identification using unmanned aerial vehicles." *Sensors* 17 (9): 2075. https://doi.org/10.3390/s17092075.

[26] Hoskere, V., J. Park, H. Yoon, and B. Spencer. 2019. “Vision-based modal survey of civil infrastructure using unmanned aerial vehicles.” *Journal of Structural Engineering.* 145 (7): 04019062. https://doi.org/10.1061/(ASCE)ST.1943-541X.0002321.

[27] Ribeiro, D., R. Santos, R. Cabral, G. Saramago, P. Montenegro, H. Carvalho, J. Correia, and R. Calcada. 2021. “Non-contact structural displacement measurement using unmanned aerial vehicles and video-based systems.” *Mechanical Systems and Signal Processing.* 160 (11): 107869. https://doi.org/10.1016/j.ymssp.2021.107869.

[28] Kalaitzakis, M. S. Kattil, N. Vitzilaios, D. Rizos, and M. Sutton. 2019. “Dynamic structural health monitoring using a DIC-enabled drone.” In Proceedings of the 2019 International Conference on Unmanned Aircraft Systems. New York: IEEE.

[29] Garg, P., R. Masimi, A. Ozdagli, S. Zhang, D. D. L. Mascarenas, M. Reda Taha, and F. Moreu. 2020. “Measuring transverse displacements using unmanned aerial systems laser doppler vibrometer (UAS-LDV): Development and field validation.” *Sensors* 20 (21): 6051. https://doi.org/10.3390/s20216051.

[30] Perry, B., and Y. Guo. 2021. “A portable three-component displacement measurement technique using an unmanned aerial vehicle (UAV) and computer vision: A proof of concept.” *Measurement* 176 (5): 109222. https://doi.org/10.1016/j.measurement.2021.109222.

[31] Nasimi, R., F. Moreu, and G. Fricke. 2023. “Sensor equipped UAS for non-contact bridge inspections: Field application.” *Sensors* 23 (1): 470. https://doi.org/10.3390/s23010470.

[32] Mondal, T. G., and M. R. Jahanshahi. 2022. “Chapter 12–Applications of depth sensing for advanced structural condition assessment in smart cities.” In *The Rise of Smart Cities,* 305-318. New York: Butterworth-Heinemann.

[33] Yoon, H., J. Shin, and B. Spencer. 2018. “Structural displacement measurement using an unmanned aerial system.” *Computer-Aided Civil and Infrastructure Engineering.* 33 (3): 183-192. https://doi.org/10.1111/mice.12338.

[34] Zhengyou Z. 1999. “Flexible camera calibration by viewing a plane from unknown orientations.” In Proceedings of the 7th IEEE International Conference on Computer Vision. New York: IEEE.

[35] Adamczyk, M., P. Liberadzki, and R. Sitnik. 2018. “Temperature compensation method for digital cameras in 2D and 3D measurement applications.” *Sensors* 18 (11): 3685. https://doi.org/10.3390/s18113685.

[36] Arun, K. S., T. S. Huang, and S. D. Blostein. 1987. “Least-squares fitting of two 3-D point sets.” *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 9 (5): 698-700. https://doi.org/10.1109/TPAMI.1987.4767965.

[37] Wu, J. 2019. “Optimal continuous unit quaternions from rotation matrices.” *Journal of Guidance, Control, and Dynamics.* 42 (4): 919-922. https://doi.org/10.2514/1.G004043.

[38] Tjahjadi, M. E., and F. Handoko. 2017. “Precise wide baseline stereo image matching for compact digital cameras.” In Proceedings of the 2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), 1-6. New York: IEEE.

[39] Harris, C., and M. Stephens. 1988. “A combined corner and edge detector.” In Proceedings of the Alvey Vision Conference. Manchester, UK: Alvey Vision Club.

[40] Bouguet, J.-Y. 2001. “Pyramidal implementation of the Affine Lucas Kanade feature tracker description of the algorithm.” *Intel Corporation* 5 (1-10) 4.

[41] Horn, B. K., and B. G. Schunck. 1981. “Determining optical flow.” *Artificial Intelligence.* 17 (1-3): 185-203. https://doi.org/10.1016/0004-3702(81)90024-2.

[42] Li, R., B. Zeng, and M. Liou. 1994. “A new three-step search algorithm for block motion estimation.” *IEEE Transactions on Circuits and Systems for Video Technology.* 4 (4): 438-442. https://doi.org/10.1109/76.313138.

[43] Sabharwal, C., and Y. Guo. 2019. “Tracking the 6-DOF flight trajectory of windborne debris using stereophotogrammetry.” *Infrastructures* 4 (4): 66. https://doi.org/10.3390/infrastructures4040066.

[44] Anwar, S., S. Khan, and N. Barnes. 2020. “A deep journey into super-resolution: A survey.” *ACM Computing Surveys.* 53 (3): 1-34. https://dl.acm.org/doi/abs/10.1145/3390462.