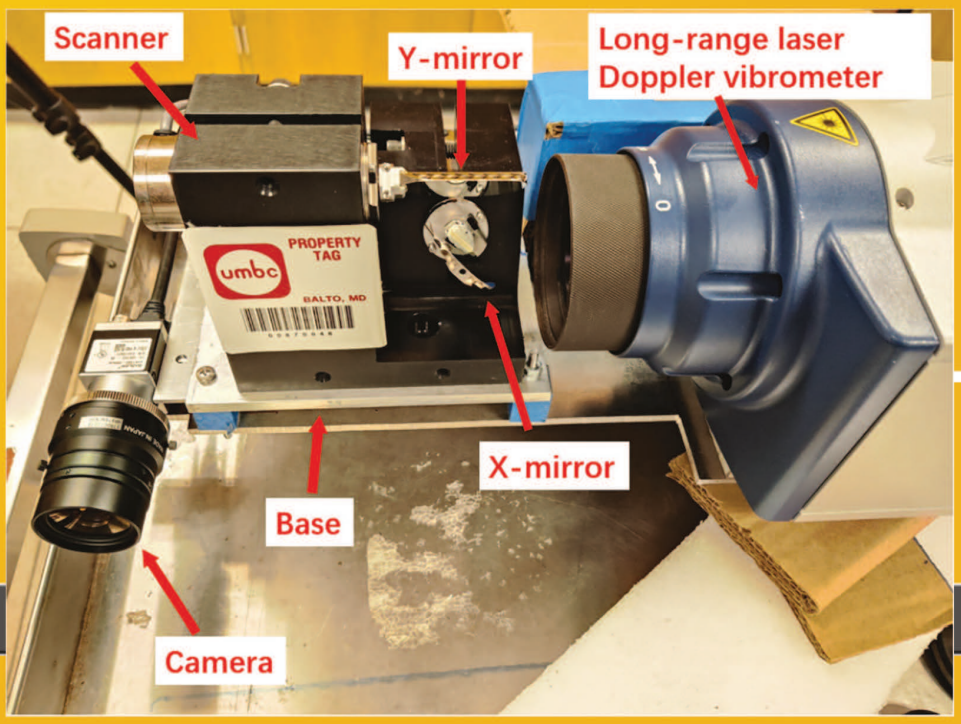
River Rapids

Conference Proceedings of the Society for Experimental Mechanics Series

Computer Vision and SDLV for Structural Dynamics, Volume 6

Janko Slavič
Daniel Rohe
Javad Baqersad



Proceedings of the 44th IMAC, A Conference and Exposition on Structural Dynamics 2026

River Publishers

Conference Proceedings of the Society for Experimental Mechanics Series

Series Editor

Kristin B. Zimmerman
Society for Experimental Mechanics, Inc.,
Bethel, USA

The Conference Proceedings of the Society for Experimental Mechanics Series presents early findings and case studies from a wide range of fundamental and applied work across the broad range of fields that comprise Experimental Mechanics. Series volumes follow the principle tracks or focus topics featured in each of the Society's two annual conferences: IMAC, A Conference and Exposition on Structural Dynamics, and the Society's Annual Conference & Exposition and will address critical areas of interest to researchers and design engineers working in all areas of Structural Dynamics, Solid Mechanics and Materials Research.

Janko Slavič · Daniel Rohe · Javad Baqersad
Editors

Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6

Proceedings of the 44th IMAC, A Conference and Exposition on Structural Dynamics 2026



River Publishers

Published, sold and distributed by:

River Publishers
Broagervej 10
9260 Gistrup
Denmark

www.riverpublishers.com

ISBN 97887-438-1445-0 (Hardback)

ISBN 97887-438-1417-7 (eBook)

https://doi.org/10.13052/rp-9788743814177

Conference Proceedings of the Society for Experimental Mechanics

An imprint of River Publishers

© The Society for Experimental Mechanics, Inc. 2026

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, or reproduction in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use. The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Preface

Computer Vision & SLDV for Structural Dynamics represent one of ten volumes of technical papers presented at the 44th IMAC, A Conference and Exposition on Structural Dynamics, organized by the Society for Experimental Mechanics, and held January 19-22, 2026. The full proceedings also include volumes on Nonlinear Structures & Systems; Dynamics of Civil Structures; Model Validation and Uncertainty Quantification; Dynamic Substructuring & Transfer Path Analysis; Special Topics & Experimental Techniques in Structural Dynamics; Dynamic Environments & Durability Testing; Topics in Modal Analysis & Parameter Identification in Structural Dynamics; Data Science in Engineering; and Structural Health Monitoring, Metamaterials, Metastructures & Additive Manufacturing and Dynamics of Biomedical & Bioinspired Systems.

This volume of proceedings shares advances in the area of computer vision, laser vibrometry, digital image correlation, photogrammetry, and optical technique and applications of these techniques for dynamic measurements.

The organizers would like to thank the authors, presenters, session organizers, and session chairs for their participation in this track.

Editors:

Janko Slavič – University of Ljubljana, Slovenia;
Daniel Rohe – Sandia National Laboratories, NM, USA;
Javad Baqersad – Kettering University

Contents

1 **Three-Dimensional Point Tracking Using UAV-Based Stereo Vision** 1
Fabio Bottalico, Christopher Niezrecki, and Alessandro Sabato

2 **Interactive AI-based Framework for Automated Damage Augmentation on Limited SHM Datasets** 9
Mohammad Mohammadzadeh and Alessandro Sabato

3 **Towards Experimental Modal Analysis of Structural Systems with High-Speed X-Ray Radiography** 17
Anthony LoRe Starleaf and David Mascareñas

4 **Mode-Shape Identification of Representative Samples with Digital Image Correlation** 27
Devin Binns, Baiza Mand, Mason Rollans, Peter Beck, Christopher Teti, Kennan Hodovic, and Jim Declerk

5 **Capturing 3D Mode Shapes with Cell Phone Videos** 43
Shawn Richardson and Mark Richardson

6 **Streaming PCA for Real-Time Structural Dynamics Tracking in Event-Based Vision** 51
Moises Felipe Silva, Anthony Starfleaf, Andre Green, Nicholas Bruns, and David Mascarenas

7 **3D Motion Detection Using Computer Vision During Earthquakes** 61
Andrés Garreaud, Ruben Boroschek, and Domingo Mery

8 **Vibration Measurement of a Rotating Structure Using an Image-Based Tracking Continuously Scanning Laser Doppler Vibrometer with a Misalignment Angle** 67
L.F. Lyu, W.D. Zhu, D. T. Will, and K. Yuan

9 **Optical Motion Magnification for Vibration-based Condition Monitoring of Hydroelectric Turbines** 85
Hussein Ali Ahmad, Marco Civera, Cecilia Surace, Christopher Niezrecki, and Alessandro Sabato

10 **Development of a Novel Robotic Continuously Scanning Laser Doppler Vibrometer for Vibration Measurement of a Beam** 95
Ke Yuan and Weidong Zhu

11 **A Brief Review of Computer Vision Technology in Pavement Management** 105
Mohamed Tareq Soliman, Feras Abla, Bilal Al-Oubaidi, Onur Avci, James Bryce, and Mahdi Nasimifar

12 **Vortex Generation by Phase-Controlled Structure-Borne Traveling Waves in a Submerged Beam: A PIV/PTV Fluid–Structure Interaction Study** 113
Rachel Cole, Samikhshak Gupta, Rohit S. Pandhare, and Vijaya V. N. Sriram Malladi

13 **Vision-Based State Estimation of a Serial Manipulator** 123
Aishani Pathak, James Halverson, Kenny DeCay Jr., Anand Iyer, Joshua Tempelman, Cole Maxwell, and Ricardo Mejia-Alvarez

14 **Thickness Estimation Using Multi-Frequency Acoustic Steady-State Excitation Spatial Spectroscopy** 137
Joseph H. Lahmann, Erica M. Jacobson, Eric B. Flynn, and Adam J. Wachtor

15 Towards Differential Laser Doppler Vibrometer Measurements using Flyable Mirrors for the Vibration Assessment of Remote Structures 147
Benjamin Halkon and Christian Rembe

16 Multi Channel Laser Radar for acousto-seimics 155
W. Zhou, O. Enriquez, G. Pandraud, E. Margallo, J. L. Rubio, and F. Beekman

17 Optimization of Flatbed UV-Printed Digital Image Correlation Speckle Patterns for Microscale Strain Mapping 161
Aastha S. Singh, Aaron P. Stebner, and Zachary D. Brunson

Chapter 1
Three-Dimensional Point Tracking Using UAV-Based Stereo Vision

Fabio Bottalico, Christopher Niezrecki, and Alessandro Sabato

Abstract Three-dimensional point tracking (3D-PT) is a powerful non-contact method for capturing displacements and deformations in structural health monitoring (SHM) applications. However, its broader adoption is limited by the need for stereo cameras to be rigidly mounted with fixed baselines and calibrated using large, cumbersome targets. These constraints make traditional 3D-PT impractical for large-scale or remote environments where access and setup are challenging. This work introduces a hybrid calibration framework that enables 3D-PT using two independently hovering unmanned aerial vehicles (UAVs), each carrying a single camera. The proposed method fuses real-time GPS and IMU data to dynamically compute extrinsic parameters, eliminating the need for fixed mounts or oversized calibration objects. To address challenges posed by UAV motion and sensor desynchronization, the approach incorporates a homography-based motion compensation technique that corrects for frame-to-frame camera movement using visual features on a planar surface. Experimental results demonstrate that the method achieves sub-centimetre tracking accuracy and a displacement error of less than 2% compared to a stationary ground-based stereo vision system, even under moderate wind disturbances. The ability to perform 3D tracking without fixed baselines or pre-installed calibration targets marks a significant step forward in making 3D-PT more scalable and field-deployable. While the current implementation requires offline compensation for sensor timing mismatches, future developments in real-time synchronization and onboard processing will further enhance its utility for dynamic, large-scale SHM applications.

Keywords Digital image correlation · stereo vision · structural health monitoring · three-dimensional point tracking · unmanned aerial vehicles

Introduction

Over recent years, computer vision techniques have gained prominence in structural health monitoring (SHM), particularly for evaluating a structure's dynamic [1]. Among these, three-dimensional digital image correlation (3D-DIC) and three-dimensional point tracking (3D-PT) have proven effective for measuring displacements in bridges [2], railroads [3], wind turbine blades [4], as well as for system identification [5] and structural dynamics studies [6].

As illustrated in Figure 1a, 3D-DIC and 3D-PT reconstruct the 3D displacement of a moving point **P** by matching its projections captured on the image planes of synchronized stereo cameras. In this work, 3D-PT is considered a subset of 3D-DIC, where a single optical target (e.g., a white circle on a dark background) is tracked instead of a stochastic speckle pattern. Before measurements can be performed, the stereo system must be calibrated to determine the intrinsic parameters (image center, focal length, and lens distortion) and the extrinsic parameters (relative orientation and distance between cameras). As shown in Figure 1b, this is typically achieved by capturing stereo images of a coded calibration object [7], such as a cross or a checkerboard with dimensions comparable to the structure under study. Because extrinsic parameters depend on the camera position, any configuration change invalidates the calibration and requires recompilation, a cumbersome process for large-scale systems. To avoid this, the cameras are mounted on a rigid bar to eliminate relative motion [1]. This requirement for scale-sensitive calibration and rigid mounting significantly limits the practicality of 3D-DIC and 3D-PT for large-scale SHM applications.

Fabio Bottalico · Christopher Niezrecki · Alessandro Sabato
Department of Mechanical and Industrial Engineering, University of Massachusetts Lowell, 1 University Avenue, Lowell, MA 01854
e-mail: Fabio_Bottalico@uml.edu; Christopher_Niezrecki@uml.edu; Alessandro_Sabato@uml.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

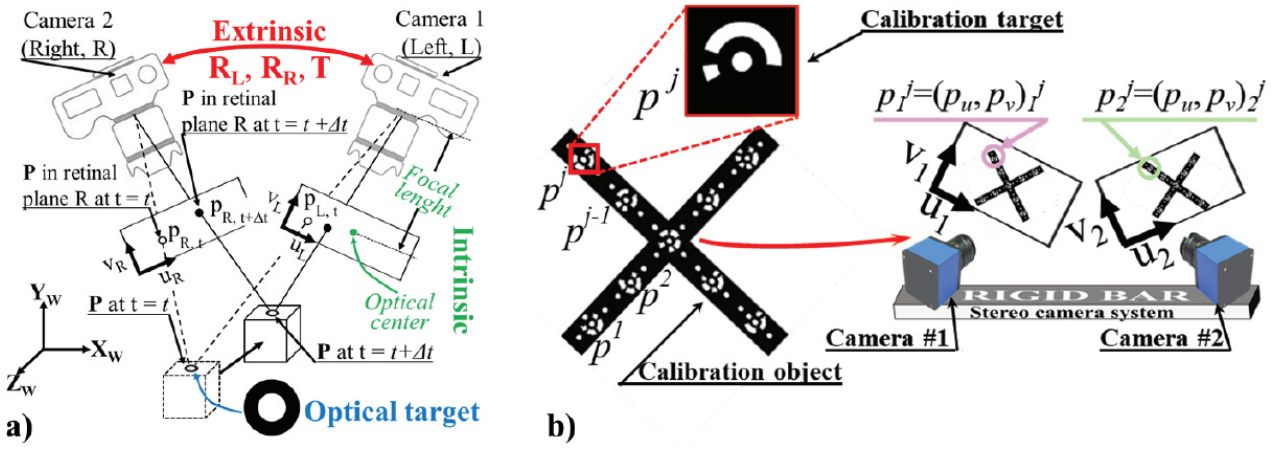


Fig. 1 Operational principles of 3D-PT: a) tracking a moving point **P** from its projection in the two retinal planes of the stereo camera system, and b) calibration using a calibration object with coded targets as seen from the two cameras.

Unmanned aerial vehicles (UAVs) have recently become powerful tools for inspecting hard-to-reach locations. UAV-based computer vision has been applied to generate 3D static maps using Structure-from-Motion (SfM) [8] and simultaneous localization and mapping (SLAM) [9], as well as to detect cracks, corrosion, and other forms of deterioration in bridges [10], dams [11], towers [12], and buildings [13] through visual and thermal imagery [14]. Beyond static mapping, UAV-mounted cameras are increasingly used for dynamic measurements. UAV oscillations during flight introduce motion artifacts, necessitating the use of fixed background points to compensate for UAV motion and recover structural displacements. This approach has been applied to experiments on trusses [15], beams [16], panels [17], and toy bridges [18], as well as in field applications on towers [19] and full-scale bridges [20]. Despite these successes, single-camera systems remain limited to 2D measurements.

To address this limitation, researchers have integrated UAVs with 3D-DIC by rigidly mounting two synchronized cameras on a single drone. Such systems have been tested on wind turbine blades [21], trusses [22], and beams [23], and deployed for crack monitoring in bridges [24]. Yet, their practicality is constrained by the short stereo baseline (typically 0.5–1 m) that a single UAV can accommodate, which restricts the field of view (FOV) to less than 2 m. To overcome these constraints, this research proposes a novel method that utilizes two independently hovering UAVs, each equipped with a single camera, to form a fully adaptable, non-fixed-baseline stereo vision system, as illustrated in Figure 2. The approach builds on prior work in sensor-based stereocamera calibration [25–27] and leverages the inertial measurement units (IMUs) and global navigation satellite systems (GNSS) integrated into commercial UAV platforms.

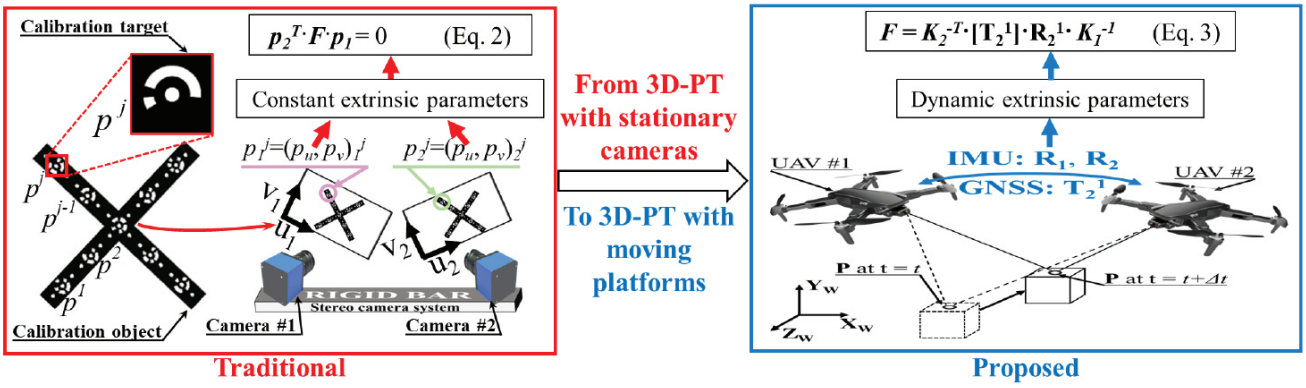


Fig. 2 Contributions of this study relative to current 3D-PT practices: extrinsic parameters are dynamically updated from UAV onboard IMU and GNSS data instead of being fixed from calibration target reprojections with rigidly mounted cameras.

This study presents the first proof-of-concept for dynamically estimating the extrinsic parameters between two cameras mounted on independent UAVs in real time. Experimental results show that the method can track the 3D displacement of

a moving target with ∼98% accuracy relative to a stationary, ground-mounted stereocamera system. The remainder of this paper is organized as follows. Section 2 introduces the mathematical framework for dynamic recalibration. Sections 3 and 4 describe the experimental setup and validation results, respectively. Section 5 concludes with key findings and directions for future research.

Methodology

The pinhole camera model forms the basis of photogrammetry and computer vision techniques. It describes how a 3D point $\boldsymbol{P} = [P_x, P_y, P_z]^T$, defined in a global reference frame *W*, is projected onto the camera's image plane as a 2D point $\boldsymbol{p} = [p_u, p_v]^T$ as expressed in Equation (1):

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

Here, the 3x3 matrix intrinsic parameters matrix $\boldsymbol{K}$ contains information about the camera's focal length (f_x, f_y), the coordinates of the optical centre (c_x, c_y), and the skew factor θ. For the applications considered in this work, $\boldsymbol{K}$ is treated as constant. The 3x4 matrix $\boldsymbol{R|T}$ describes the extrinsic parameters and consists of the nine components of a rotation matrix $\boldsymbol{R}$ and the translation vector $\boldsymbol{T}$. Calibration is traditionally performed by solving Equation (1) for $\boldsymbol{K}$ and $\boldsymbol{R|T}$ using multiple images of a calibration target with known geometry (i.e., the coordinates of all points $\boldsymbol{p}^j$ shown in Figure 1b). When a 3D point $\boldsymbol{p}^j$ is projected onto the two image planes, its representation appears as points $\boldsymbol{p}_1{}^j$ and $\boldsymbol{p}_2{}^j$ in the two cameras, connected by a relationship expressed as:

$$\boldsymbol{p_2^T}\ \boldsymbol{F}\ \boldsymbol{p_1} = 0 \tag{2}$$

Here, **F** is a 3x3 singular matrix of rank two, known as the fundamental matrix, which includes the intrinsic and extrinsic parameters of the stereo camera as:

$$\boldsymbol{F} = \boldsymbol{K}_2^{-T}\ \left[\boldsymbol{T_2^1}\right]_x \boldsymbol{R_2^1}\ \boldsymbol{K}_1^{-1} \tag{3}$$

In Equation (3), $\boldsymbol{K}_1{}^{-1}$ is the inverse of the intrinsic parameters' matrix of Camera 1 and $\boldsymbol{K}_2{}^{-T}$ is the transpose of the intrinsic parameter matrix of Camera 2. $[\boldsymbol{T}_2{}^1]_x$ is the cross-product matrix of the translation vector $\boldsymbol{T}_2{}^1$ (i.e., the relative distance between the cameras), while $\boldsymbol{R}_2{}^1$ is the relative rotation matrix of Camera 2 expressed in Camera 1's reference frame. Traditional target-based calibration methods compute $\boldsymbol{F}$ directly from Equation (2) by identifying multiple pairs of corresponding points $(\boldsymbol{p}_1, \boldsymbol{p}_2)_i$ in stereo images of a calibration object. The intrinsic and extrinsic parameters are then extracted from the decomposition of $\boldsymbol{F}$ using Equation (2). By contrast, sensor-based calibration methods bypass this step and rely directly on Equation (3), where $\mathbf{R}_2{}^1$ and $\mathbf{T}_2{}^1$ are measured from the UAVs' GNSS and IMU sensors (see Figure 2), as defined in Equations (4) and (5).

$$\begin{cases} T_1^W(t) = N * lla2ned\left(T_1^{WGS84}(t) - T_1^{WGS84}(0)\right) \\ T_2^W(t) = N * lla2ned\left(T_2^{WGS84}(t) - T_1^{WGS84}(0)\right) \end{cases} \tag{4}$$

$$\boldsymbol{R}_i^W = \begin{bmatrix} cos(\alpha_i) & -sin(\alpha_i) & 0 \\ sin(\alpha_i) & cos(\alpha_i) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} cos(\beta_i) & 0 & sin(\beta_i) \\ 0 & 1 & 0 \\ -sin(\beta_i) & 0 & cos(\beta_i) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\gamma_i) & -sin(\gamma_i) \\ 0 & sin(\gamma_i) & cos(\gamma_i) \end{bmatrix} \tag{5}$$

In Equation (4), the translation vectors $\mathbf{T}_i{}^W$ of the two cameras are obtained from the GNSS sensors, which are reported in WGS84 latitude–longitude–altitude (LLA) coordinates and then converted into the North-East-Down (NED) reference frame using MATLAB's *lla2ned()* function. In Equation (5), α, β, and γ are the roll, yaw, and pitch angles measured by the IMUs on the UAVs, following the NED convention.

In practice, however, the GNSS, IMU, and camera sensors on commercial UAVs typically operate at different or insufficient sampling rates, which complicates the data processing. For instance, high-frequency oscillations of a hovering UAV are not captured by the GNSS, and GNSS measurements are often not perfectly synchronized with the camera. These limitations introduce errors in the estimation of extrinsic parameters, leading to inaccuracies in the triangulated positions of tracked points. To mitigate these issues, the images themselves must be exploited to refine the estimation of extrinsic parameters. A practical solution is motion correction via homography, as described in Equation (6).

$$\boldsymbol{h_2} = \boldsymbol{H}\boldsymbol{h_1} \tag{6}$$

Here, $\boldsymbol{H}$ is a non-singular 3x3 homogeneous matrix, and $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ are the points in the image selected as a reference. Although similar in form, Equation (6) imposes different constraints than Equation (2). By matching feature points between an image acquired at $t = 0$ and another at $t = t_i$, Equation (6) can be solved for $\boldsymbol{H}$. The resulting homography is then fused with the IMU–GNSS data to compensate for the relative motion of the two UAVs. In summary, the calibration procedure involves two steps: first, estimating the extrinsic parameters of the stereo system using IMU and GNSS sensors; and second, applying the homography method to correct any residual motion between the cameras.

Experimental Setup

Two tests were conducted to evaluate the accuracy of the proposed method for tracking the 3D displacement of an object using two cameras mounted on independent UAVs. The setup employed two DJI Matrice 300 RTK UAVs, each equipped with a gimbal-mounted camera. UAV 1 carried a Sony A7R-V with a 35 mm fixed focal length lens (UAV-Camera 1), while UAV 2 carried a DJI Zenmuse XT2 with an 8 mm fixed focal length lens. The intrinsic parameters of both cameras were calibrated beforehand, and because the lenses are fixed, the intrinsic matrices $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ remained constant across all measurements.

Test 1 evaluated the feasibility of estimating the extrinsic parameters of a stereovision system by combining real-time kinematic global navigation satellite system (RTK-GNSS) data for the translation vector $\boldsymbol{T}$ with orientation data from the high-sensitivity IMUs embedded in the UAVs' gimbals for the rotation matrix $\boldsymbol{R}$. For this experiment, the UAVs were positioned as shown in Figure 3a, forming a stereovision system with an approximate baseline of 8 m and a working distance of 15 m. The UAV-mounted cameras captured stereo image sequences of a calibration checkerboard while simultaneously recording GNSS and IMU data. The extrinsic parameters derived from the UAV sensors were then compared with those obtained from image-based calibration. Test 1 was designed to minimize synchronization-related errors and to assess the geometric performance of the proposed approach without interference from timing mismatches among the GNSS, IMU, and cameras.

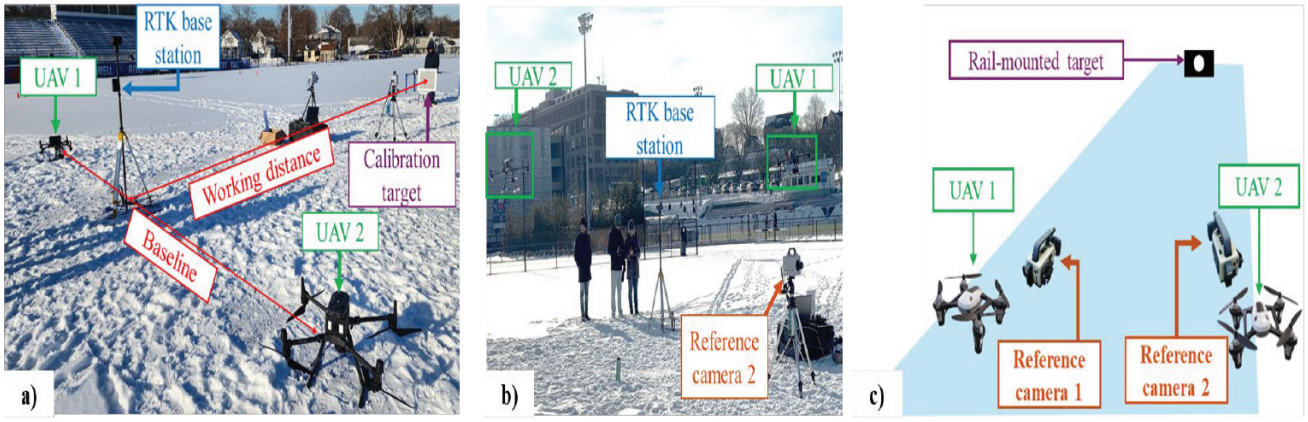


Fig. 3 Experimental setup: a) Test 1, showing the two UAVs, the RTK base station, and the calibration checkerboard used for image-based calibration, b) UAVs hovering behind the ground-based cameras during Test 2, and c) schematic (not to scale) of the experimental setup for Test 2.

Test 2 evaluated the feasibility of applying the proposed calibration method to track a moving object using two cameras mounted on independently hovering UAVs. The same UAVs from Test 1 were flown in manual mode, each maintaining an altitude of ∼2.5 m. As shown in Figures 3b and 3c, the UAVs were positioned behind two ground-based cameras that served as a reference system. A rail-mounted sliding carriage was installed on supports, with a high-contrast optical target attached to the slider. The target was translated from right to left in 12 increments. At each position, both the ground-based stereocamera system and the UAV-mounted cameras captured five images. During acquisition, wind gusts of up to 8 $m \cdot s^1$ caused significant UAV oscillations. The RTK-GNSS and IMU data collected from the two UAVs were again used to compute the extrinsic parameters of the stereovision system formed by UAV 1 and UAV 2.

Results

The results of **Test 1** are summarized in Table 1, comparing the extrinsic parameters estimated with the GNSS–IMU system against those from conventional image-based calibration. As shown in Table 1, the extrinsic parameters obtained with the proposed calibration method closely match those from traditional image-based calibration, particularly in the angular measurements. The yaw angle β (i.e., the separation angle of the two cameras in the horizontal plane through their optical centres) differs by only 0.012°, corresponding to a relative error of 0.037%, demonstrating excellent agreement. The pitch α and roll γ angles from the proposed method are nearly zero, reflecting both the parallel alignment of the cameras to the ground and the effectiveness of the gimbal stabilization. By contrast, the image-based calibration produced nonzero pitch and roll values, likely due to the calibration object being too small relative to the field of view (a well-known scale sensitivity issue that this work aims to overcome).

Table 1 Extrinsic parameters of UAV–Camera 2 relative to UAV–Camera 1 from image-based and GNSS–IMU calibration in Test 1.

	α (°)	β (°)	γ (°)	T_X (*mm*)	T_Y (*mm*)	T_Z (*mm*)
Image-based	2.275	32.249	-0.940	-8012.50	34.53	2009.60
Proposed	0.022	32.261	0.007	-8116.30	304.79	2041.20
Difference	2.253	-0.012	-0.947	103.80	-270.26	-31.60

The translation components also show strong consistency. Differences in the horizontal T_X and vertical T_Z components of the translation vector correspond to errors of 1.29% and 1.57%, respectively. The only significant discrepancy appears in the lateral component T_Y, where the methods differ by 270.26 mm. This deviation stems from the altitude quantization resolution of the RTK-GNSS system (∼304.79 mm, or 1 ft) used on the commercial UAV platform. Crucially, this reflects a hardware constraint rather than a limitation of the proposed method.

Figure 4 presents the results of tracking the optical target (see Figure 3c) with two cameras mounted on independent UAVs. Two cases are shown: *1)* extrinsic parameters estimated using only GNSS–IMU data, and *2)* extrinsic parameters

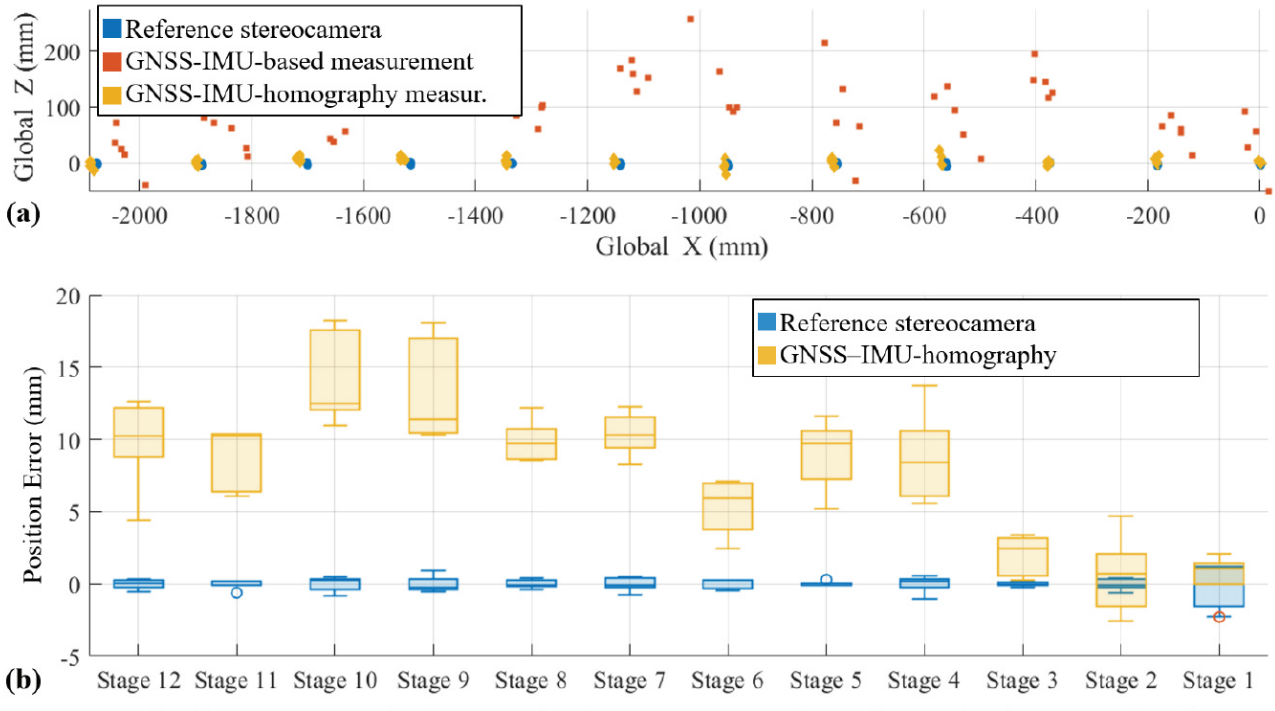


Fig. 4 Displacement of the optical target: a) locations before and after motion compensation, showing improved alignment, and b) whisker plot of deviations from the ground-based reference.

refined using GNSS–IMU data with homography-based motion compensation. The compensation approach significantly improves accuracy, reducing errors to the centimetre-level (Figure 4a), as the measured displacement converges more closely to the ground-based stereocamera reference. Without compensation, GNSS–IMU measurements deviate from the reference by up to $\pm$50 mm in mean displacement, with standard deviations of 50 to 100 mm. In contrast, GNSS–IMU–homography results closely match the reference values (golden dots compared to red dots in Figure 4a).

The whisker plot in Figure 4b zooms in on the results shown in Figure 4a and further confirms this improvement: the maximum absolute mean displacement error decreases to just 15 mm, compared to the $\pm$50 mm range observed without compensation. At Stage 10, where the total displacement is 1700 mm, this corresponds to a relative error of only 0.88%. At Stage 12, with a total displacement of 2090 mm, the 10 mm error yields an even smaller relative error of 0.48%. In addition to improved accuracy, the precision of the UAV-based measurements increased substantially. The standard deviation of displacement dropped from 50-100 mm to 5-10 mm, representing an order-of-magnitude enhancement. These results clearly demonstrate the effectiveness of the proposed motion compensation method in mitigating UAV-induced errors during 3D tracking.

Conclusion

A novel calibration method is presented for 3D motion tracking using two independently hovering UAVs, each carrying a camera. Unlike conventional 3D-PT, which requires rigidly mounted cameras and large calibration targets, the proposed approach dynamically estimates extrinsic parameters by combining RTK-GNSS and IMU data and applies homography to compensate for UAV-induced instabilities. This is the first demonstration of stereo vision calibration and 3D-PT performed with two moving UAVs rather than fixed cameras.

Two field tests validated the method. In Test 1, the extrinsic parameters estimated using the proposed GNSS–IMU approach closely matched those obtained from traditional image-based calibration, with angular deviations of nearly zero and translation errors of less than 2%. In Test 2, a moving optical target was tracked using the UAV-mounted cameras. While raw GNSS and IMU data introduced errors due to sensor desynchronization, incorporating homography-based motion compensation reduced the error to below 1% and improved precision by an order of magnitude.

Overall, the proposed method achieves accuracy comparable to conventional calibration techniques while offering greater flexibility and scalability. By removing the need for rigid bars and calibration panels, it enables 3D tracking in scenarios previously impractical for 3D-PT. Although the current accuracy does not yet meet the millimeter-level demands for SHM, this study establishes a new pathway for UAV-based 3D-PT. Future work should prioritize improving sensor synchronization and integrating higher-rate sensors to enhance reliability and achieve the precision required for full-scale SHM deployment.

Acknowledgments This material is based upon work supported by the National Science Foundation (NSF) under Grant Number 2440348 - *CAREER: Enhancing Measurements of Dynamic Features in Large-Scale Structures via Three-Dimensional Aerial Stereovision*. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. B. Pan, "Digital image correlation for surface deformation measurement: historical developments, recent advances and future goals," Measurement Science and Technology, vol. 29, no. 8, p. 082001, 2018.
2. L. Ngeljaratan and M. A. Moustafa, "Structural health monitoring and seismic response assessment of bridge structures using target-tracking digital image correlation," Engineering Structures, vol. 213, p. 110551, 2020.
3. A. Sabato and C. Niezrecki, "Feasibility of Digital Image Correlation for railroad tie inspection and ballast support assessment," Measurement, vol. 103, pp. 93-105, 2017.
4. P. Poozesh, A. Sabato, A. Sarrafi, C. Niezrecki, P. Avitabile, and R. Yarala, "Multicamera measurement system to evaluate the dynamic response of utility-scale wind turbine blades," Wind Energy, vol. 23, no. 7, pp. 1619-1639, 2020.
5. M. Shafiei Dizaji, M. Alipour, and D. K. Harris, "Leveraging Full-Field Measurement from 3D Digital Image Correlation for Structural Identification," Experimental Mechanics, vol. 58, pp. 1049-1066, 2018.
6. D. Gorjup, J. Slavič, and M. Boltežar, "Frequency domain triangulation for full-field 3D operating-deflection-shape identification," Mechanical Systems and Signal Processing, vol. 133, 2019.
7. Z. Zhang, "A flexible new technique for camera calibration," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pp. 1330-1334, 2000.
8. S. Jiang, C. Jiang, and W. Jiang, "Efficient structure from motion for large-scale UAV images: A review and a comparison of SfM tools," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 167, pp. 230-251, 2020.

9. A. Gupta and X. Fernando, “Simultaneous Localization and Mapping (SLAM) and Data Fusion in Unmanned Aerial Vehicles: Recent Advances and Challenges,” Drones, vol. 6, no. 4, p. 85, 2022.
10. S. Feroz and S. Abu Dabous, “UAV-Based Remote Sensing Applications for Bridge Condition Assessment,” Remote Sensing, vol. 13, no. 9, p. 1809, 2021.
11. S. Zhao, F. Kang, J. Li, and C. Ma, “Structural health monitoring and inspection of dams based on UAV photogrammetry with image 3D reconstruction,” Automation in Construction, vol. 130, p. 103832, 2021.
12. H. Wang, H. Zhou, H. Liu, Z. Huang, and M. Feng, "Research on Determining the Inspection Point of Multirotor UAV Power Tower," Mathematical Problems in Engineering, vol. 2021, no. 1, p. 8894055, 2021.
13. T. Rakha and A. Gorodetsky, “Review of Unmanned Aerial System (UAS) applications in the built environment: Towards automated building inspection procedures using drones,” Automation in Construction, vol. 93, pp. 252-264, 2018.
14. S. Dabetwar, N. N. Kulkarni, M. Angelosanti, C. Niezrecki, and A. Sabato, “Sensitivity analysis of unmanned aerial vehicle-borne 3D point cloud reconstruction from infrared images,” Journal of Building Engineering, vol. 58, p. 105070, 2022.
15. G. Chen, Q. Liang, W. Zhong, X. Gao, and F. Cui, “Homography-based measurement of bridge vibration using UAV and DIC method,” Measurement, vol. 170, p. 108683, 2021.
16. G. Liu, J. Luo, Q. Yang, S.-s. Law, and C. He, “Non-contact Structural Displacement Measurement using UAV and DIC with Non-coplanar Fixed Region of Interest,” Measurement, vol. 242, p. 115936, 2025.
17. Y. Weng, J. Shan, Z. Lu, X. Lu, and B. F. Spencer Jr., “Homography-based structural displacement measurement for large structures using unmanned aerial vehicles,” Computer-Aided Civil and Infrastructure Engineering, vol. 36, no. 9, pp. 1114-1128, 2021.
18. Z. Wu, G. Chen, Q. Ding, B. Yuan, and X. Yang, “Three-Dimensional Reconstruction-Based Vibration Measurement of Bridge Model Using UAVs,” Applied Sciences, vol. 11, no. 11, p. 5111, 2021.
19. H. Habeenzu, P. McGetrick, S. Taylor, and D. Hester, “UAS-based bridge displacement measurement using two cameras with non-overlapping fields of view,” Automation in Construction, vol. 167, p. 105687, 2024.
20. Y. Han, G. Wu, and D. Feng, “Vision-based displacement measurement using an unmanned aerial vehicle,” Structural Control and Health Monitoring, vol. 29, no. 10, p. e3025, 2022.
21. A. Khadka, B. Fick, A. Afshar, M. Tavakoli, and J. Baqersad, “Non-contact vibration monitoring of rotating wind turbines using a semi-autonomous UAV,” Mechanical Systems and Signal Processing, vol. 138, p. 106446, 2020.
22. G. Lavezzi, M. Ciarcià, K. Won, and M. Tazarv, “A DIC-UAV based displacement measurement technique for bridge field testing,” Engineering Structures, vol. 308, p. 117951, 2024.
23. M. Kalaitzakis, N. Vitzilaios, D. C. Rizos, and M. A. Sutton, “Drone-Based StereoDIC: System Development, Experimental Validation and Infrastructure Application,” Experimental Mechanics, vol. 61, no. 6, pp. 981-996, 2021.
24. D. Reagan, A. Sabato, and C. Niezrecki, “Feasibility of using digital image correlation for unmanned aerial vehicle structural health monitoring of bridges,” Structural Health Monitoring, vol. 17, no. 5, pp. 1056-1072, 2018.
25. A. Sabato, N. A. Valente, and C. Niezrecki, “Development of a Camera Localization System for Three-Dimensional Digital Image Correlation Camera Triangulation,” IEEE Sensors Journal, vol. 20, no. 19, pp. 11518-11526, 2020.
26. F. Bottalico, C. Niezrecki, K. Jerath, Y. Luo, and A. Sabato, “Sensor-Based Calibration of Camera’s Extrinsic Parameters for Stereophotogrammetry,” IEEE Sensors Journal, vol. 23, no. 7, pp. 7776-7785, 2023.
27. Bottalico, F., Valente, N. A., Niezrecki, C., Jerath, K., Luo, Y., & Sabato, A. (2025). Rapid 3D Camera Calibration for Large-Scale Structural Monitoring. Remote Sensing, 17(15), 2720.

Chapter 2
Interactive AI-based Framework for Automated Damage Augmentation on Limited SHM Datasets

Mohammad Mohammadzadeh and Alessandro Sabato

Abstract Structural Health Monitoring (SHM) plays a crucial role in ensuring the reliability and operational efficiency of large-scale engineering structures such as wind turbines. While extensive visual inspection datasets exist, they predominantly contain images of healthy wind turbine blades, significantly restricting the available data for supervised machine learning damage detection. Conventional data augmentation techniques typically include and increase irrelevant information and features, such as background and sky artifacts, thus reducing the detection accuracy and effectiveness of the machine learning model. To address these challenges, an interactive framework is introduced for damage augmentation, designed explicitly for limited SHM datasets. In this framework, users can upload available data and explicitly annotate damage using bounding boxes, enabling the automated extraction and characterization of damage such as cracks or erosion. Leveraging advanced generative artificial intelligence (AI) techniques, the framework synthesizes realistic and diverse damage instances by using Stable Diffusion fine-tuned with low-rank adapters. In this research, the proposed method is validated by augmenting real-world wind turbine blade imagery, demonstrating improved diversity and realism of the synthesized damages. Results highlight the method's potential to effectively enhance training datasets, paving the way for better-performing supervised models. This research contributes a practical and scalable augmentation solution, addressing critical data scarcity issues in SHM, and providing a robust foundation for future AI-driven damage detection systems.

Keywords Generative Artificial Intelligence · Structural Health Monitoring · Data Augmentation · Machine Learning · Wind Turbine Blades

Introduction and related work

The long service life of wind turbines depends on effective inspection and maintenance practices. Manual blade inspection is slow, costly, and risky. Missed or late detection of surface cracks can lead to downtime and structural failures. Therefore, automated visual detection is desirable for safer operations, lower costs per inspection, and improved maintenance scheduling [1].

Training reliable damage detectors is difficult with small labeled datasets. Data collection at scale is expensive, and many field images contain no cracks, creating severe class imbalance and limited coverage of crack appearances. Conventional augmentation strategies, such as rotation, rescaling, and flipping, improve feature invariance but cannot generate new crack morphologies or realistic surface representations [2]. More sophisticated approaches, including image-mixing and copy–paste approaches [3], extend data diversity but still fall short of bridging the gap between limited data and high-dimensional models. Recent augmentation pipelines have emerged that leverage generative models, such as Contrastive Language–Image Pretraining (CLIP) and Stable Diffusion (SD) [4, 5], to synthesize realistic cracks on real backgrounds. Generative Adversarial Network (GAN) pipelines have become a common choice, but they often require larger datasets and careful training to avoid model collapse and texture artifacts [6]. In contrast, diffusion models can be adapted with fewer parameters [7]. Low-Rank Adaptation (LoRA) makes this adaptation efficient and more stable in the small-data setting used here [8].

In this research, we introduce an interactive augmentation framework designed to address the problem of data scarcity in structural health monitoring of wind-turbine blades. The approach focuses on generating realistic crack appearances that

Mohammad Mohammadzadeh · Alessandro Sabato
Department of Mechanical and Industrial Engineering, University of Massachusetts Lowell, 1 University Avenue, Lowell, MA 01854
e-mail: Mo_moha@uml.edu; Alessandro_Sabato@uml.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

resemble those observed in practice, thereby enriching small datasets that would otherwise limit the reliability of supervised machine learning models. By combining user guidance with generative AI, the framework ensures that synthetic data remain consistent with the application domain and can be seamlessly integrated into training pipelines. This enables the creation of diverse, realistic datasets that improve model performance and help overcome one of the key barriers to deploying AI-based crack detection in real-world settings.

This work contributes: i) a practical, end-to-end, human-in-the-loop framework that couples LoRA-tuned diffusion with mask-constrained placement for blade cracks; ii) a direct comparison of alpha compositing and Poisson blending for inserting fine synthetic cracks on the surfaces of WTBs; and iii) a controlled, small-data evaluation across the five scenarios investigated in this study demonstrating how generative augmentation can substitute for real training data.

Methods

A data-efficient augmentation pipeline was developed to improve the detection of surface cracks on WTBs. Real crack regions were first extracted as 512×512 patches from annotated blade images. A SD v1.5 UNet model was then fine-tuned with low-rank adapters (LoRA) using these patches. Two ranks were trained, with rank 8 retained for its superior visual realism. During augmentation, a *per-sample loop* was used: one synthetic crack patch was generated, screened for realism in-app, and—if accepted—placed within user-defined feasible regions of a real background image. Two blending mechanisms were employed: i) alpha blending after thresholding to isolate crack pixels, and ii) gradient-domain Poisson blending via seamless cloning. Each placement accepted by the user produces a single bounding-box annotation. Any overlap with existing boxes triggered a resampling of the location. Synthetic data were restricted to the training split, while validation was kept real-only with a single "crack" class. Five training scenarios were constructed: baseline, adding 200 alpha-blended synthetic cracks, adding 200 Poisson-blended synthetic cracks, adding 200 Poisson-blended synthetic cracks on crack-free (CF) images plus original data, and 200 synthetic-only. YOLOv8m was trained with the provided configuration, with the best checkpoint selected automatically. The results of the five scenarios were evaluated in terms of precision, recall, and mean Average Precision (mAP) at 0.5 Intersection over Union (IoU).

LoRA fine-tuning of Stable Diffusion 1.5

SD v1.5 [9] was fine-tuned on curated 512×512 crack patches extracted from high-resolution images of WTBs. The goal was to train the model on the visual concept of a "hairline surface crack on a white blade". In SD, images are represented in a compressed latent space and are synthesized by iteratively denoising random noise under text guidance. For fine-tuning, each patch was paired with a single caption, and the text encoder remained the same except for the final portion of the training. Images were resized with Lanczos interpolation and normalized to the [-1, 1] range. Training was performed with batch size of 4 for 1,200 steps and a learning rate of 10^{-4}, with other settings left at library defaults. To adapt the model efficiently, LoRA [10] was applied to the cross-attention layers of the UNet. Rather than updating all weights, LoRA introduces small, task-specific updates to each affected linear layer, expressed as:

$$\Delta W = \frac{\alpha}{r} AB \tag{1}$$

Where ΔW is the low-rank update applied to the frozen weight matrix, A and B are trainable thin matrices of rank r, and α is a scaling factor that controls the contribution of the update. The effective weight is thus $W^{'}$=W+ΔW, with the base weights W kept frozen. The LoRA scaling was set to $\alpha = r$, giving the update unit gain. Two ranks were tested (i.e., $r = 4$ and $r = 8$), with $r = 8$ retained for producing crack appearances that were more realistic and closer to the training patches during in-app screening. Fine-tuning followed the library's standard diffusion training, where, at a random timestep t, Gaussian noise is added to the latent space representation z_t. The UNet is then trained to predict this noise by minimizing the mean-squared error (MSE) loss function:

$$\mathrm{L_{MSE}} = \mathrm{E}[\|\epsilon - \widehat{\epsilon}_\theta(\mathrm{z_t}, \mathrm{t}, \mathrm{c})\|_2^2] \tag{2}$$

where z_t denotes the noisy latent variable, c the text conditioning that guides the generation, ϵ the known injected noise, and $\widehat{\epsilon}_\theta(z_t, t, c)$the predicted noise from the UNet parameterized by θ. This loss measures the squared difference between the injected noise and the network's prediction, ensuring that the model learns to gradually denoise the latent representation in alignment with the conditioning. In practice, only the low-rank LoRA parameters are adjusted, guiding the denoising process to generate crack-like features consistent with the supervised patches while leaving the base model largely unchanged.

Per-Sample Generation, Screening, and Placement

Augmentation followed a one-by-one loop. A synthetic 512×512 crack patch was generated with the LoRA-tuned SD model and screened in-app for realism. Non-crack-like outputs were discarded, while accepted patches were placed within user-defined feasible regions on real background images from the 100-image training split. The 180-image validation split was kept real-only (no synthetic crack was added).

Feasible regions for placing synthetic cracks were defined as polygonal masks, saved as binary images $M \in \{0,1\}^{H \times W}$, where H and W denote the image height and width. In these masks, a pixel value of $M(x,y) = 1$ indicates that the pixel at location (x,y) is a valid placement area, while $M(x,y) = 0$ indicates that placement is not allowed. The complete set of admissible coordinates is therefore:

$$\Omega = \{(x,y)\,|\,M\,(x,\ y) = 1\}, \quad |\Omega| = \sum_{x=1}^{W}\sum_{y=1}^{H} M\,(x,y)\,. \tag{3}$$

Here, Ω is the set of all valid pixel locations, and $|\Omega|$ is its cardinality (the total number of allowed pixels). A placement coordinate (u,v), representing the pixel location of the patch center on the background image, was then drawn according to a uniform probability distribution $p((u,v))$:

$$p\,((u,v)) = \begin{cases} |\Omega|^{-1},\ (u,v)\ \in\ \Omega \\ 0,\ \ otherwise \end{cases} \tag{4}$$

In other words, each valid pixel in Ω has equal probability of being selected. The chosen coordinate (u,v) was used as the center of the synthetic crack patch. No rotation or rescaling was applied, and no minimum distance from the image edges was enforced. Before including the newly generated synthetic image, a human-in-the-loop visual check was performed to ensure that the new bounding box did not improperly overlap existing annotations and that the placement looked plausible on the surface. If an issue was observed, a new $(u,\ v)$ was sampled and the placement was redone. For each accepted placement, background image ID, random seed, mask identifier, $(u,\ v)$ coordinates, selected blending method, and resulting YOLO bounding box were logged.

Blending Method

Two placement mechanisms were employed. The first used alpha compositing, where only crack pixels from the generated patch were transferred to the background image. For each generated patch, a grayscale version was computed, and an interactive per-sample threshold was applied (without morphological operations) to produce a binary mask $\alpha(x,y) \in \{0,1\}$. Let $P(x,y)$ denote the generated crack patch (RGB), $B(x,y)$ the background image (RGB), and $R(x,y)$ the resulting composite image. The compositing operation is given by

$$R\,(x,y) =\ \alpha\,(x,\ y)\ P\,(x,y) +\ \ (1-\alpha\,(x,y))\ B\,(x,y)\,, \tag{5}$$

The second mechanism used gradient-domain blending (Poisson blending) via OpenCV's *seamlessClone* function with the NORMAL_CLONE option. In this case, a binary mask was derived from the crack patch by thresholding its grayscale at 240, then refined with a 3×3 morphological close–open sequence. Let Ω denote the masked region and $\partial\Omega$ its boundary. The blended image I inside Ω was obtained by solving the discrete Poisson objective with boundary conditions taken from the background B:

$$\min \sum_{(i,\ j)\in\Omega} \|\nabla \mathrm{I}\ (\mathrm{i},\ \mathrm{j}) -\ \nabla \mathrm{P}(\mathrm{i},\ \mathrm{j})\|_2^2\ s.t. I\,(i,j) = B\,(i,\ j)\ for\,(i,\ j)\ \in \partial\Omega \tag{6}$$

This optimization preserves the crack gradients from the patch P while ensuring the color and illumination at the boundary match the background B. In practice, the cloning center was set to the placement coordinates (u,v). If the placement was near the image border, masks and patches were cropped, and the OpenCV solver returned the final blended result R.

Experimental Design

Five training scenarios were evaluated to isolate how synthesis and blending on detection. In all cases, synthetic data was only used for training purposes, while validation (180 images) remained real-only. The baseline training split contained 100 real images. Across all scenarios, the same augmentation loop, mask-constrained placement, overlap-redo policy, label format, and YOLOv8m training configuration was applied. Only the composition of the training set varied.

- *Baseline (Exp. 1).* The original 100 real images were used without any synthetic additions, providing the reference detector performance.
- *Alpha-200 on train (Exp. 2).* 200 synthetic cracks were added to the 100 real images of Exp. 1 using the alpha composite method described in (2.6) and used for training. Cracks were generated and screened one at a time, then placed within feasible regions of a training image. Multiple cracks per image were allowed until 200 total placements were reached. Placement centers were sampled uniformly over the valid mask of the chosen background, and any visually problematic overlap with existing boxes triggered a resampling of the location. No rotation or scaling was applied, and the alpha masks were formed by per-sample thresholding without morphology (see Section 2.3).
- *Poisson-200 on train (Exp. 3).* 200 synthetic cracks were added to the 100 real training images using the Poisson blending formulation shown in (2.7) via seamlessClone (NORMAL_CLONE). As in Exp. 2, multiple cracks per image were allowed until 200 accepted placements were reached. Cloning centers, mask construction, and border handling followed the approach discussed in Section 2.3.
- *Poisson-200 on CF + original (Exp. 4).* A separate set of 200 crack-free images was each populated with one Poisson-blended synthetic crack. These images were then combined with the original 100 real training images. Feasible-region masks were enforced on the crack-free images as in the original training images.
- *Synthetic-only-200 (Exp 5).* A set of 200 crack-free images, each populated with one Poisson-blended synthetic crack, was constructed. No original real training images were used. This scenario tests the lower bound achievable when training relies solely on synthetic data, but evaluation remains on real images.

Table 1 Details of the five experiments considered in the study.

Scenario	Real train images	Synthetic cracks	Blending	Per-image rule
Exp. 1	100	0	-	-
Exp. 2	100	200 on train	Alpha (2.6)	Multiple
Exp. 3	100	200 on train	Poisson (2.7)	Multiple
Exp. 4	100	200 on CF + original	Poisson (2.7)	One
Exp. 5	0	200 on CF	Poisson (2.7)	One

All detectors were trained with Ultralytics YOLOv8m (pretrained yolov8m.pt) for 100 epochs using 1024×1024 images, a batch size of 16, and SGD with a base learning rate of 0.001. Checkpoints were selected by the trainer's best metric. While the training dataset differs by scenario, the validation and test sets were fixed and real-only.

Evaluation

Evaluation was performed on the held-out real validation set. Each predicted box $\hat{b}$ was greedily matched to at most one ground-truth box b based on the highest Intersection over Union (IoU). A match was accepted only if

$$IoU\left(\hat{b}, b\right) = \frac{area(\hat{b} \cap b)}{area(\hat{b} \cup b)} \geq 0.5 \tag{7}$$

With this rule, matched predictions were counted as true positives (TP), unmatched predictions as false positives (FP), and unmatched ground-truth boxes as false negatives (FN). Precision and recall at a given confidence threshold were defined as:

$$Precision = \frac{TP}{TP + FP}, \tag{8}$$

$$Recall = \frac{TP}{TP + FN}. \tag{9}$$

Predictions sorted by confidence and threshold were swept to form the precision—recall curve at the fixed IoU threshold of 0.5. Average Precision (AP) at 0.5 IoU, denoted $AP_{0.5}$ was computed as the area under this curve using the discretization provided by the library:

$$AP_{0.5} = \int_0^1 p\left(r\right)\, dr \;\approx \sum_{k=1}^{K} p_k \left(r_k - \, r_{k-1}\right), \tag{10}$$

where p_k and r_k are the precision and recall values at the *k-th* operating point. Since only one class was used, the mean Average Precision at 0.5 IoU equals $AP_{0.5}$,

Results

Data Summary and Patch Body

Training used 100 real images with 154 annotated cracks, while validation used 180 real images with 234 annotated cracks. All reported metrics were computed on the validation split to test robustness under a broader distribution of crack instances while keeping the training dataset deliberately small. For the real images, 154 curated 512×512 crack patches were extracted to adapt SD v1.5 with LoRA. Figure 1 shows the visual basis for synthesis and the effect of the LoRA rank. Figure 1(a) displays a 4×4 grid of real training patches, while Figures 1(b) and (c) display 4×4 grids of synthetic patches generated with LoRA rank 4 and rank 8, respectively, using identical prompts and sampler settings. Rank 8 was retained because its samples more closely matched the real training patches in shape, contrast, and background cleanliness during in-app screening. Figure 2 illustrates the effect of the two blending mechanisms. Alpha compositing transfers only the crack pixels after thresholding, which sometimes leaves faint edges around the patch. In contrast, Poisson blending better integrates the synthetic cracks with the background surface, producing smoother transitions in shading and illumination.

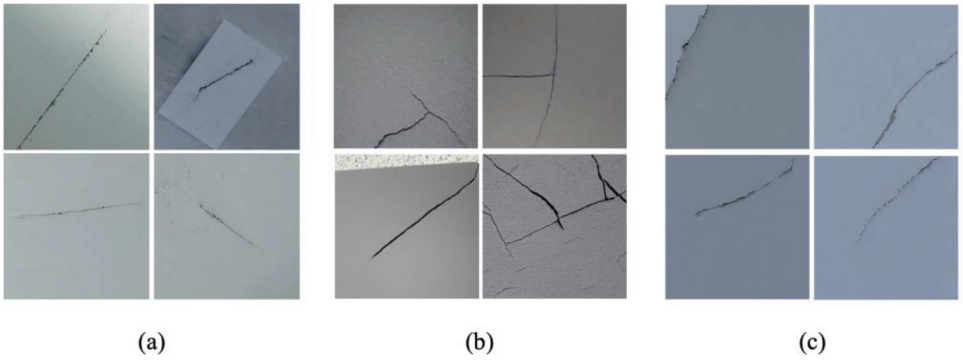


Fig. 1 Patch grids. (a) Real 4×4 training patches. (b) Synthetic 4×4 patches from SD+LoRA rank 4. (c) Synthetic 4×4 patches from SD+LoRA rank 8 (selected for subsequent experiments).

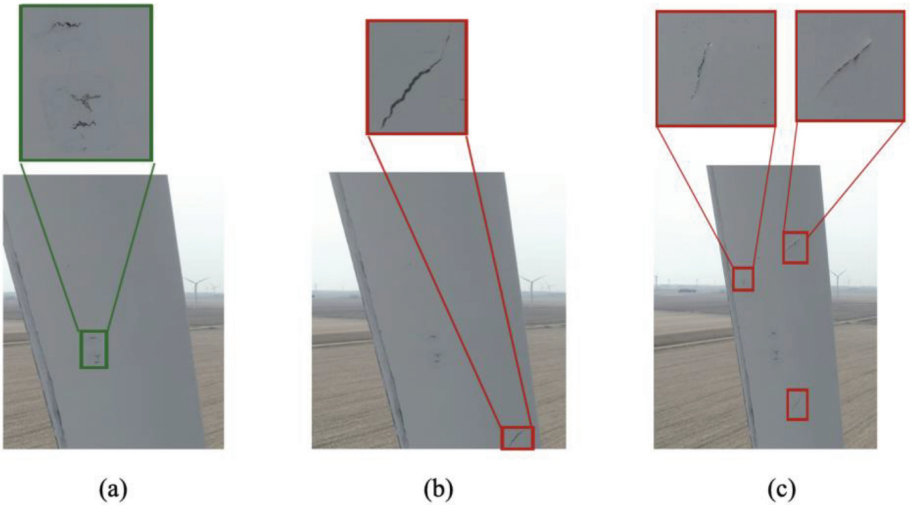


Fig. 2 Side-by-side comparison of blending methods. (a) Original training image with real cracks. (b) Example with alpha-blended synthetic cracks. (c) Example with Poisson-blended synthetic cracks. (The boxes highlight example cracks: green for real and red for synthetic. The specific synthetic instances differ across panels).

End-to-end app snapshot

The augmentation was carried out in an end-to-end app that integrates generation, placement, and auto-annotation into a single workflow (Figure 3). The LoRA-adapted diffusion model produces one crack at a time, which is screened in-app for realism. If accepted, a training image is selected, a feasible region is chosen, and the crack is placed at a sampled location. Alpha or Poisson blending is applied, and a YOLO-format box is written automatically. If a placement overlaps an existing box or looks implausible, a new location is sampled immediately.

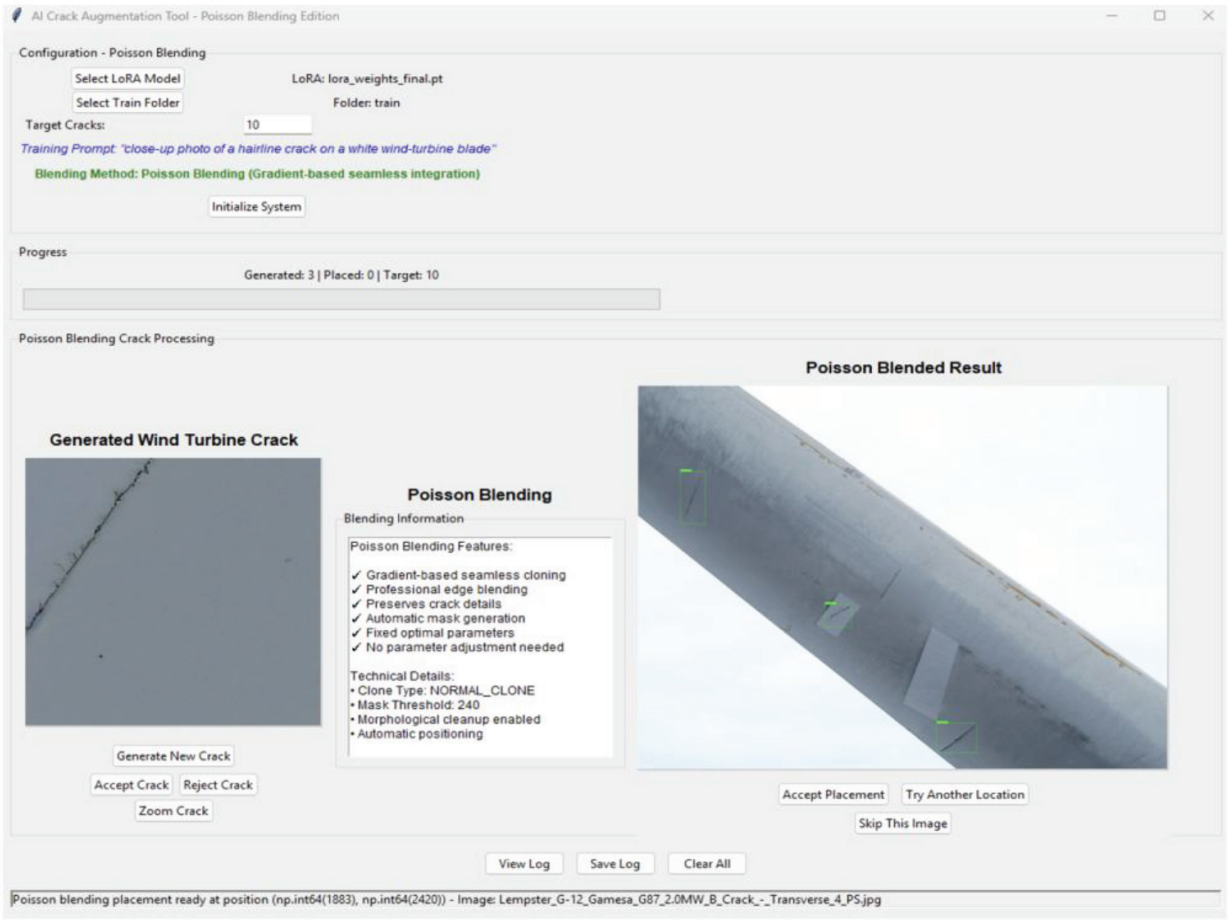


Fig. 3 App snapshot. The interface supports per-sample crack generation and screening, mask-constrained placement and automatic YOLO annotation and logging—all within one window.

Quantitative Performance

All metrics identified in Section 2.5 were computed on the validation split (i.e., 180 images, 234 cracks). The same YOLOv8m configuration was used across all scenarios studied and only the composition of the training dataset was different (i.e., 100 images, 154 cracks).

Table 2 Comparison of experimental scenarios on detection metrics.

Scenario	mAP	Recall	Precision
Exp. 1 Baseline	48.8	46.6	61.6
Exp. 2 Apha-200 on train	46.8	45.3	53.5
Exp. 3 Poisson-200 on train	**59.4**	**55.5**	**70.9**
Exp. 4 Poisson-200 on CF + original	52.4	53.6	65.0
Exp. 5 Synthetic-only-200	30.3	33.3	42.5

The Poisson-200 on-train scenario *(Exp. 3)* achieved the best results with mAP = 59.4, Recall = 55.5, and Precision = 70.9. Relative to the baseline *(Exp. 1)*, absolute gains of +10.6 in mAP, +8.9 in Recall, and +9.3 in Precision were observed. The Poisson-200 crack-free + original scenario *(Exp. 4)* also improved recall and precision over baseline (53.6 and 65, respectively) with a smaller mAP increase (54.4), suggesting that inserting synthetic images into the original training dataset is most effective under the given constraints. The Alpha-200 on-train scenario *(Exp. 2)* underperformed baseline across all metrics (i.e., 46.8 / 45.3 / 53.5), consistent with qualitative artifacts observed with alpha compositing on varying backgrounds.

The synthetic-only scenario *(Exp. 5)* was included to quantify how much of the baseline can be recovered using *only* synthetic data. With 200 Poisson-blended images, it reached 30.3 / 33.3 / 42.5. This corresponds to about 62% of baseline mAP (30.3 vs. 48.8), 72% of baseline recall (33.3 vs. 46.6), and 69% of baseline precision (42.5 vs. 61.6). Thus, while a small synthetic body of synthetic images alone was insufficient to match the real-data baseline, it recovered a substantial fraction of the baseline performance under the same evaluation on real images, establishing a practical lower bound for zero-real training in this setting. Figure 4 shows sample of predicted cracks on the wind turbine blade.



Fig. 4 Samples of predicted cracks on wind turbine using YOLOv8.

Conclusion and Future Work

A data-efficient, end-to-end augmentation framework and app for surface crack detection on wind turbine blades is presented in this work. SD v1.5 was adapted with LoRA (rank 8) using curated 512×512 crack patches. Augmentation proceeded one sample at a time with in-app realism screening, mask-constrained placement on real training images, and automatic YOLO annotations. Two blending strategies were tested: alpha compositing, which transfers crack pixels after thresholding, and Poisson blending, which preserves shading and illumination on white gel-coat surfaces. Results of the tests performed showed how Poisson blending delivered the most reliable improvements over baseline, while the synthetic-only case, though lowest, recovered a substantial fraction of baseline performance, establishing a practical lower bound for zero-real training.

This work establishes an end-to-end, human-in-the-loop app that unifies generation, placement, blending, and annotation into a YOLO-ready workflow. Moreover, it is the first controlled comparison of alpha and Poisson blending for fine crack insertion under strict data limits; and an evaluation of synthetic–real mixing strategies, showing that Poisson-based insertions on real training images yield the largest gains. Future work will automate realism scoring, introduce distribution-aware placement, extend labels to segmentation masks, broaden detector baselines and metrics, and ablate LoRA hyperparameters. This study shows that generative augmentation, paired with human oversight, can significantly improve crack detection in low-data situations.

References

1. Ciang, C. C., Lee, J. R., & Bang, H. J. (2008). Structural health monitoring for a wind turbine system: a review of damage detection methods. *Measurement science and technology*, *19*(12), 122001.
2. Sengupta, P., Mehta, A., & Rana, P. S. (2023, July). Enhancing performance of deep learning models with a novel data augmentation approach. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-7). IEEE.

3. Ghiasi, G., Cui, Y., Srinivas, A., Qian, R., Lin, T. Y., Cubuk, E. D., ... & Zoph, B. (2021). Simple copy-paste is a strong data augmentation method for instance segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 2918-2928).
4. Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125*, *1*(2), 3.
5. Song, J., Meng, C., & Ermon, S. (2020). Denoising diffusion implicit models. *arXiv preprint arXiv:2010.02502*.
6. Luleci, F., Catbas, F. N., & Avci, O. (2023). Generative adversarial networks for labeled acceleration data augmentation for structural damage detection. *Journal of Civil Structural Health Monitoring*, *13*(1), 181-198.
7. Fang, H., Han, B., Zhang, S., Zhou, S., Hu, C., & Ye, W. M. (2024). Data augmentation for object detection via controllable diffusion models. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision* (pp. 1257-1266).
8. Zhang, Z., Zhou, Y., & Ma, J. (2025). Defect-LoRA: Controllable defect data augmentation based on low-rank adaptation for surface defect recognition under limited data. *Expert Systems with Applications*, *268*, 126251.
9. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10684-10695).
10. Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., ... & Chen, W. (2022). Lora: Low-rank adaptation of large language models. *ICLR*, *1*(2), 3.

Chapter 3
Towards Experimental Modal Analysis of Structural Systems with High-Speed X-Ray Radiography

Anthony LoRe Starleaf and David Mascareñas

Abstract This work introduces a novel data analysis technique for performing 3D modal analysis of structures with optically inaccessible internal geometry using radiographic data acquired via high-speed X-ray imaging systems. We reconstruct the static configuration of the structure by employing Neural Radiance Field (NeRF)-inspired inverse rendering, resulting in a volumetric representation defined by a three-dimensional grid of attenuation coefficients. To characterize the structure's response, we employ an isoparametric finite-element (FE) framework. The core innovation lies in directly learning deflection vectors for each time step using a fixed-point iterative grid warping algorithm, which estimates nodal deflections from dynamic radiographs. By comparing reconstructed radiographs against actual captured frames, the deflection estimates are refined until they implicitly represent the true response of the structure (up to the fidelity of the FE mesh). The final deflection estimates serve as inputs to a simple Rayleigh quotient–based modal analysis procedure, enabling accurate extraction of modal parameters. It should be noted that the estimation of modal parameters is entirely decoupled from the displacement measurement algorithm. There is a wide range of literature which addresses the problem of obtaining modal parameters from FE response data; in principle, any of these methods could be used with our approach. We selected the Rayleigh quotient due to its simplicity and straightforward implementation. Finally, the proposed method offers significant advantages for non-contact, high-resolution modal analysis. By considering X-ray radiography, we overcome a key limitation of traditional vision-based modal analysis – the requirement that the structure's surfaces must be externally visible.

Introduction

Experimental modal analysis (EMA) is a widely used technique for updating models of civil, mechanical, and aerospace structures [1]. Typically, EMA is performed by applying contact sensors, such as accelerometers and strain gauges, to the surface of a structure, then exciting the structure with a known force and recording its response. While contact sensors are well-studied and reliable, they come with a few limitations. First, when applied to thin or light (low mass) structures, the sensors themselves significantly impact the dynamics. Second, they can be expensive and time-consuming to apply. Third, they only provide pointwise measurements of the response, necessitating a large number of them to characterize the behavior of the structure.

In response to the limitations of contact sensors, several vision-based approaches to EMA have been proposed. Phase-based motion magnification and mode shape identification [2, 3] is a technique that extracts mode shapes and frequencies from video footage of vibrating structures. The authors of [4] proposed an x-ray-based digital volume correlation (DVC) method for extracting motion from stereo x-ray images. Recent work from Sandia National Laboratory [5] proposed the QuintX system, which is a 5-view stereo x-ray imaging system designed for EMA. Inspired by this work, we seek to develop a computational method of extracting full-field deflection measurements of a vibrating structure from dynamic, multi-view x-ray images.

Anthony LoRe Starleaf · David Mascareñas
e-mail: aglor@lanl.gov; dmascarenas@lanl.gov

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

Methods

Our methodology is broken into three major steps: (1) static volume reconstruction, (2) deflection field characterization, and (3) system identification. In the first two steps, we represent the attenuation field associated with the object of interest by a 3D grid of attenuation values (Figure 2). In the static reconstruction phase, we use the Adam optimizer [6] to iteratively update the 3D attenuation grid until radiographs rendered by our differentiable x-ray rendering engine match those captured experimentally. This results in a volumetric representation of the attenuation field associated with the structure of interest which implicitly encodes its geometry. In the deflection field characterization phase, we represent structural deformations using nodal displacement vectors on a finite element (FE) mesh. Using a differentiable grid warping procedure, we deform the static attenuation field according to the deflection field specified by the nodal displacement vectors and render the new volume. Again, the Adam optimizer is used to find a set of nodal displacements that minimizes the difference between the rendered and experimentally captured set of x-ray images at each time-step. This results in a set of nodal deflections for each time-step that can be input into a system identification algorithm. A flowchart describing this process is shown below (Figure 1).

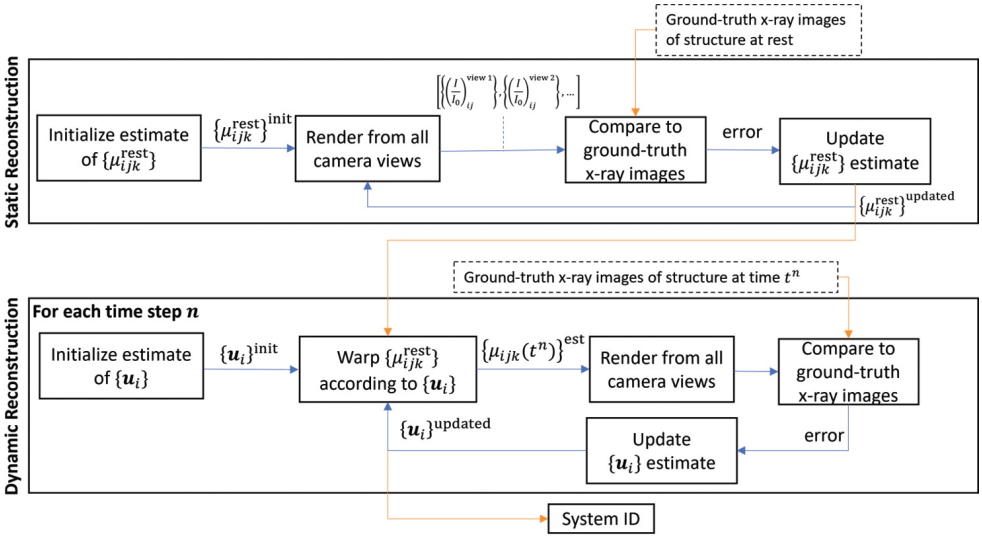


Fig. 1 Flowchart describing our methodology. Blue arrows represent the flow of information within a single step of our procedure. Orange arrows represent the flow of information between steps. Gradients are not propagated along orange arrows. The remainder of this section discusses each component of the process in more detail.

3D Voxel Grid Representation

We represent the structure's attenuation field using a 3D grid of attenuation coefficients μ_{ijk}, interpolated according to the standard trilinear interpolation formula for a uniform grid given by

$$\overline{\mu}\,(x,y,z) = \sum_{ijk} w_{ijk}\,(x,y,z)\,\mu_{ijk}$$

$$w_{ijk}\,(x,y,z) = \frac{1}{\Delta x \Delta y \Delta z} \begin{cases} |x-x_{ijk}| \cdot |y-y_{ijk}| \cdot |z-z_{ijk}| & \text{if } \left\|\left(\boldsymbol{x}-\boldsymbol{x}_{ijk}\right)/\Delta \boldsymbol{x}\right\|_{\infty} \leq 1 \\ 0 \text{ otherwise} \end{cases}$$

where $\|\cdot\|_{\infty}$ denotes the supremum norm. A visualization of this interpolation procedure is given below in Figure 2. The grid is represented in 2D for the sake of visual clarity.

Applying this interpolation formula allows a continuous function $\mu : \mathbb{R}^3 \to \mathbb{R}$ to be represented by a discrete set of points $\{\mu_{ijk}\}_{0,0,0}^{N_x,N_y,N_z}$. Similarly to Figure 2, Figure 3 below shows a 2D analog of this procedure.

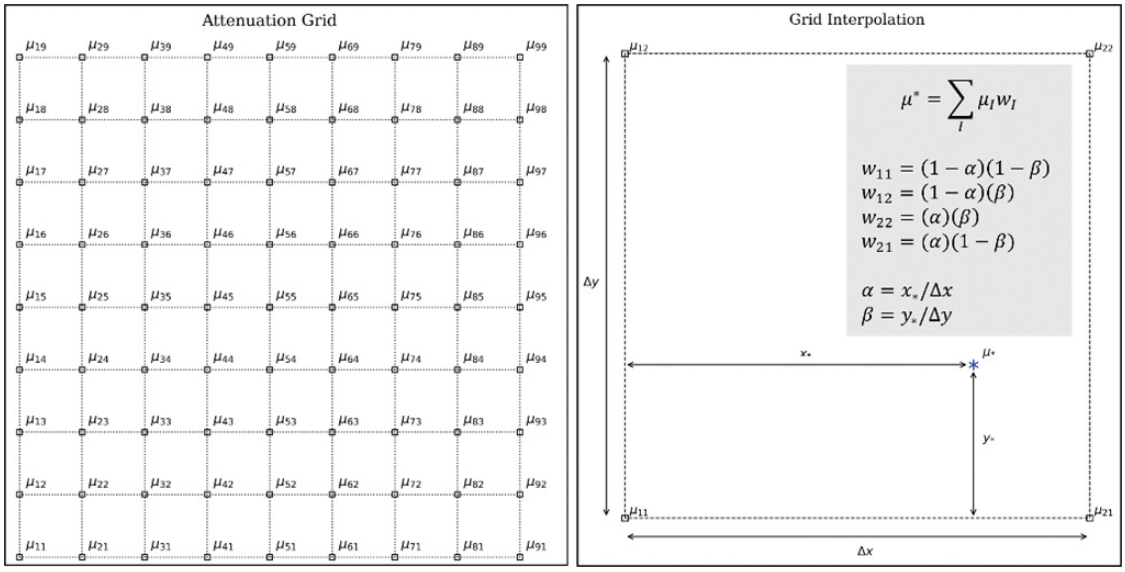


Fig. 2 Left: 2D representation of the 3D voxel grid. Right: bilinear interpolation (2D analog of trilinear interpolation) formula. In 3D, the grid would have an additional axis extending into/out of the page.

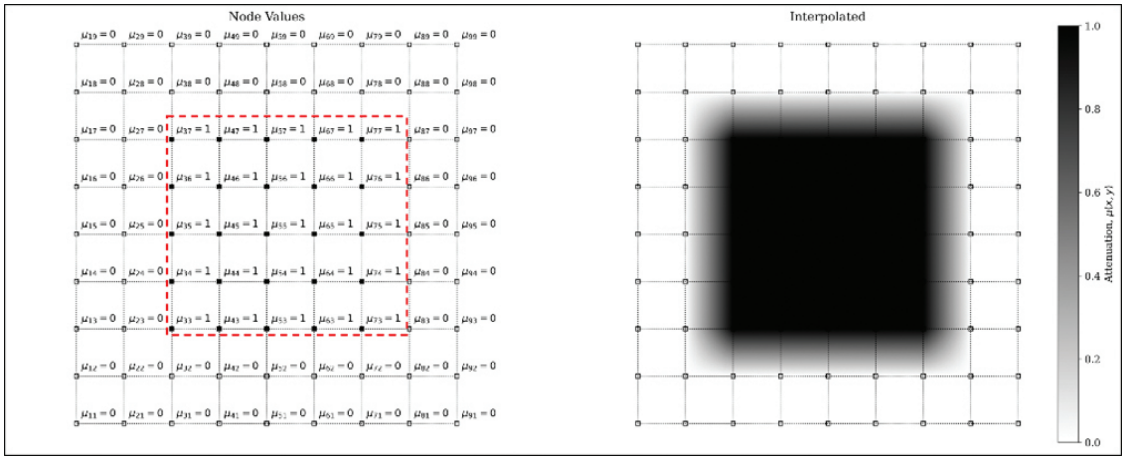


Fig. 3 Left: 2D voxel grid; The red square encloses nodes whose attenuation values are set to 1. Right: The interpolated continuous attenuation field $\mu(x, y)$.

Differentiable X-Ray Rendering

Since we use gradient-based optimization to perform both the static reconstruction and deflection field characterization, it is important for the x-ray rendering procedure to be differentiable. Fortunately, our parameterization of the attenuation field is continuous with respect to the node values μ_{ijk}. To render the attenuation field, we simulate an isotropic x-ray source casting rays through the voxel grid and striking a scintillator plate with a photodiode array behind it. For simplicity, we do not differentiate between the photodiode array and scintillator plate, and assume that the scintillator plate itself has pixels whose intensities can be recorded to form an image. A 2D analog of the rendering configuration is shown in Figure 4.

In this configuration, x-rays are emitted by the source at a uniform intensity I_0 and are attenuated by the object (represented by the 3D voxel grid). The rays then strike a scintillator plate/pixel array with received intensity I. The relationship between I, I_0, and μ_{ijk} is governed by Beer's Law.

$$\log \frac{I}{I_0} = -\int \mu\left(x, y, z; \mu_{ijk}\right) ds$$

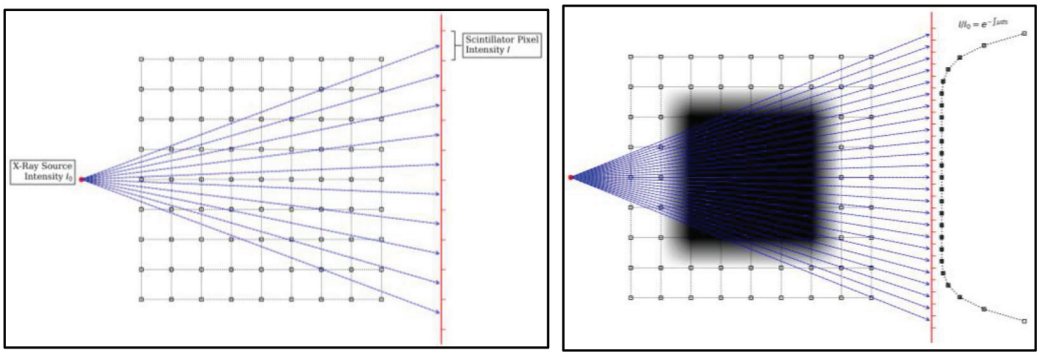


Fig. 4 Left: 2D analog of the rendering configuration. In 3D, the scintillator plate would be a bounded 2D plane. Right: Example of a 1D radiograph rendered from a 2D attenuation grid. The rightmost graph shows received intensity I/I_0 as a function of pixel position.

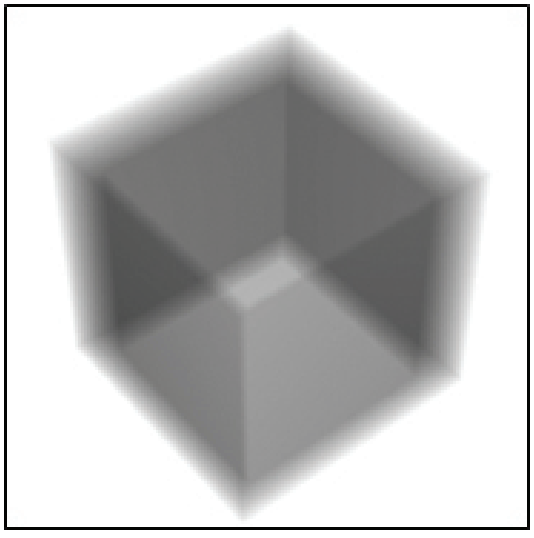


Fig. 5 Example of a 2D radiograph rendered from a 3D attenuation field.

where ds indicates that the integral is carried out lengthwise along each ray. Applying this procedure to a 3D volume would yield a 2D array of values $\left(\frac{I}{I_0}\right)_{ij}$ where ij indexes pixels. An example of such an image is shown in Figure 5. This is a simplification of the method of x-ray rendering employed in [7]. As is evident from the lack of dependency on wavelength in the equations above, we assume monochromatic x-ray generation.

Differentiable Grid Warping

The concept of "grid warping" may at first seem rather unintuitive. Its purpose in our methodology is to allow the statically reconstructed attenuation field representing the structure's rest configuration to be deformed according to estimates of the structure's deflection field at a given time t. Since the attenuation field is represented by a 3D voxel grid, the most straightforward way to represent the new configuration of the structure at time t is to construct a new voxel grid μ^t_{ijk} such that

$$\text{Interp}\left(\mu^t_{ijk}, \begin{bmatrix} x + u^t(x,y,z) & y + v^t(x,y,z) & z + w^t(x,y,z) \end{bmatrix}\right) = \text{Interp}(\mu_{ijk}, \begin{bmatrix} x & y & z \end{bmatrix})$$

where 'Interp' represents the interpolation operator. The central challenge in accomplishing this is determining the original position (i.e. where the node was located before the deformation was applied) of each node in the new grid μ^t_{ijk}. Once the original position $(x_{ijk}, y_{ijk}, z_{ijk})$ of each node in the new grid has been determined, the nodal attenuation values μ^t_{ijk} can be obtained by interpolating in the original grid at $(x_{ijk}, y_{ijk}, z_{ijk})$. Figure 6 demonstrates this idea in 2D.

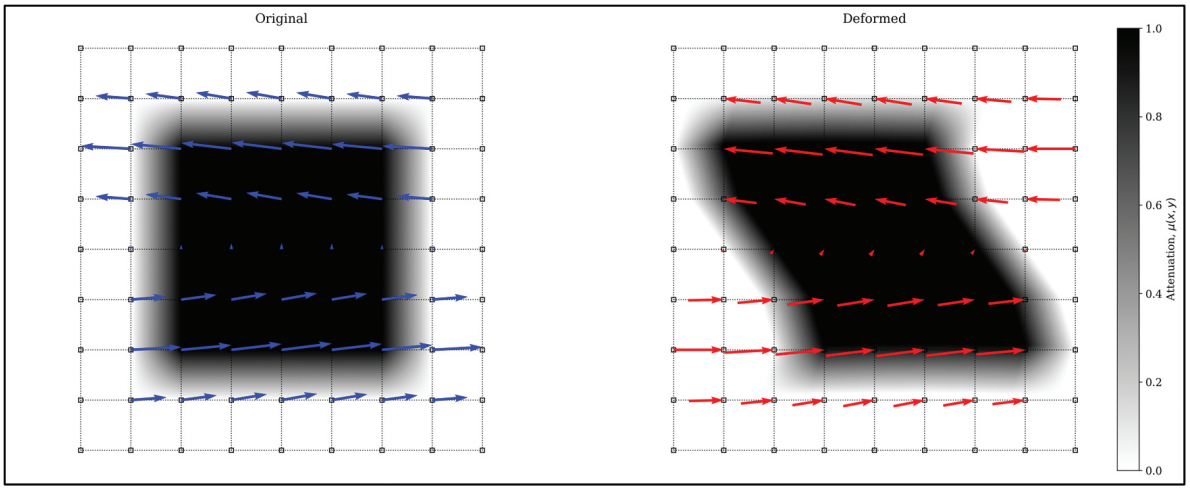


Fig. 6 Left: statically reconstructed grid μ_{ij} with blue arrows defining a forward deformation field for the object. Right: new grid μ_{ij}^t at time t representing the object after deformation. Red arrows point from the original position of each grid node to its current position.

In order to determine the original positions (x, y, z) of points in deformed space $(X, Y, Z) = (x, y, z) + (u^t, v^t, w^t)$, we apply an α-damped fixed-point iteration [8] with the mapping

$$g(\boldsymbol{x}) = (1-\alpha)\,\boldsymbol{x} + \alpha\,(\boldsymbol{X} - \boldsymbol{\delta}(\boldsymbol{x}))$$

$$\boldsymbol{x}_{\text{est}}^{k+1} = g(\boldsymbol{x}_{\text{est}}^{k})$$

where $\boldsymbol{x} = (x, y, z)$, $\boldsymbol{X} = (x + u^t(\boldsymbol{x}), y + v^t(\boldsymbol{x}), z + w^t(\boldsymbol{x}))$, and $\boldsymbol{\delta}(\boldsymbol{x}) = (u^t(\boldsymbol{x}),\ v^t(\boldsymbol{x}),\ w^t(\boldsymbol{x}))$. This is guaranteed to converge as long as g is a contraction on $\mathbb{R}^3$. That is, as long as g satisfies

$$\|g(\boldsymbol{x}) - g(\boldsymbol{y})\|_2 \leq \|\boldsymbol{x} - \boldsymbol{y}\|_2 \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$$

In the $\alpha = 0$ case, $g(\boldsymbol{x}) = \boldsymbol{x}$ and convergence is guaranteed. However, this does not incorporate any knowledge of the deformation field $\boldsymbol{\delta}(\boldsymbol{x})$ and will "converge" to the wrong solution unless $\boldsymbol{x}_{\text{est}}^0 = \boldsymbol{x}$. In practice, we initialize $\boldsymbol{x}_{\text{est}}^0 = \boldsymbol{X}$, so this only works if $\boldsymbol{\delta}(\boldsymbol{x}) = \boldsymbol{0}$. In the $\alpha = 1$ case, the iteration only converges if $\boldsymbol{\delta}(\boldsymbol{x})$ is a contraction on $\mathbb{R}^3$, which is usually the case. In our implementation, we initialize α to 1 and halve it when $\boldsymbol{x}_{\text{est}}^{k+1}$ stops improving over $\boldsymbol{x}_{\text{est}}^{k}$. Importantly, this grid warping procedure is differentiable with respect to $\boldsymbol{\delta}(\boldsymbol{x})$.

We use PyTorch's backward-mode auto-differentiation to perform all gradient-based optimizations discussed in this paper. Backward-mode auto-differentiation stores intermediate results of all computations to be used for computing derivatives. This means that the memory usage associated with our procedure grows linearly with the number of fixed-point iterations needed to deform the grid. Empirically, we notice that the fixed-point iteration procedure tends to converge in approx. 10-20 iterations. This is not enough to cause a problem on most mid-grade consumer GPUs (i.e. the NVIDIA GeForce RTX 3070 Ti Laptop GPU) but may cause issues if the method is applied with very fine voxel grids or highly irregular deflection fields.

System Identification

Once the nodal displacement vector $\boldsymbol{u}(t)$ of the FE mesh has been estimated for each time step, we use a version of the Rayleigh quotient to obtain estimates of the quantity E/ρ. For completeness, we provide a short derivation of the quotient below. Begin with the discrete equations of motion for the structure in free vibration.

$$\boldsymbol{M}\ddot{\boldsymbol{u}}(t) + \boldsymbol{K}\boldsymbol{u}(t) = \boldsymbol{0}$$

Assuming material homogeneity gives $\boldsymbol{M} = \rho\widehat{\boldsymbol{M}}$ and $\boldsymbol{K} = E\widehat{\boldsymbol{K}}$. Rewriting,

$$\widehat{\boldsymbol{M}}\ddot{\boldsymbol{u}}(t) + \frac{E}{\rho}\widehat{\boldsymbol{K}}\boldsymbol{u}(t) = \boldsymbol{0}$$

Applying the Fourier transform and denoting the Fourier transform of $\boldsymbol{u}(t)$ by $\boldsymbol{\phi}(\omega)$,

$$\frac{E}{\rho}\widehat{\boldsymbol{K}}\boldsymbol{\phi}\left(\omega\right) - \omega^2\widehat{\boldsymbol{M}}\boldsymbol{\phi}\left(\omega\right) = \mathbf{0}$$

Moving the second term to the right and left multiplying by $\boldsymbol{\phi}^T$,

$$\frac{E}{\rho}\boldsymbol{\phi}^T\widehat{\boldsymbol{K}}\boldsymbol{\phi} = \omega^2\boldsymbol{\phi}^T\widehat{\boldsymbol{M}}\boldsymbol{\phi}$$

Finally, this may be rearranged to obtain

$$\frac{E}{\rho} = \frac{\boldsymbol{\phi}^T\widehat{\boldsymbol{M}}\boldsymbol{\phi}}{\boldsymbol{\phi}^T\widehat{\boldsymbol{K}}\boldsymbol{\phi}}\omega^2$$

In practice, we obtain the modal frequencies ω_i from the time-series values of $q_i(t)$ estimated in the dynamic reconstruction phase. We then estimate E/ρ according to the following square-amplitude-weighted average.

$$\frac{E}{\rho} = \frac{\sum_i \left(\frac{\boldsymbol{\phi}_i^T\widehat{\boldsymbol{M}}\boldsymbol{\phi}_i}{\boldsymbol{\phi}_i^T\widehat{\boldsymbol{K}}\boldsymbol{\phi}_i}\omega_i^2\right)\left(\max_t q_i^2\left(t\right)\right)}{\sum_i \max_t q_i^2\left(t\right)}$$

It should be noted that once the nodal deflections have been determined in the dynamic reconstruction step, any method of system identification may be used. We only use this one because it is simple.

Results & Discussion

We applied the above procedure to a synthetic experiment involving a hollow 3D box in free vibration. For the sake of expediency in evaluating this approach, an excessive number of imager views and samples were used. In the future we intend to repeat this experiment with fewer views and a lower sampling rate. Table 1 defines numerical values for several of the parameters which define our experiment.

Table 1 Numerical values of experimental parameters.

Experiment Parameter	Value	Units
Number of Imagers	32	-
Sampling Rate	3.33	kHz
Elastic Modulus, E	70	GPa
Density, ρ	2700	$\mathrm{kg{\cdot}m^3}$
Poisson's Ratio, ν	0.3	-
Structure Length/Width/Height	1.0 / 1.0 / 1.0	m
Voxel Grid Size	[128, 128, 128]	-
Voxel Grid Resolution	10.94	mm

Synthetic Data Generation

To validate our procedure, we applied it to a simulated hollow box in free vibration. A hollow box was selected because it is a simple structure, and the cavity in the middle allows us to leverage the x-ray imagers' ability to see internal surfaces. The box was initially deformed, then released and allowed to vibrate freely. Figure 7 shows the box in its rest configuration, as well as its shape at $t = 0$.

The box was meshed with linear tetrahedral (TET4) elements, and its motion was simulated according to the standard damping-free FE equation $\boldsymbol{M}\ddot{\boldsymbol{u}}\left(t\right) + \boldsymbol{K}\boldsymbol{u}\left(t\right) = \mathbf{0}$ using an implicit Newmark-Beta scheme with $\Delta t = 3 \times 10^{-4}$. Figure 8 shows the configuration of x-ray imagers around the box.

The rendering procedure described in Section II-B was used to generate simulated x-ray images of the box in its rest configuration (Figure 8), and in its deformed configuration at each time step.

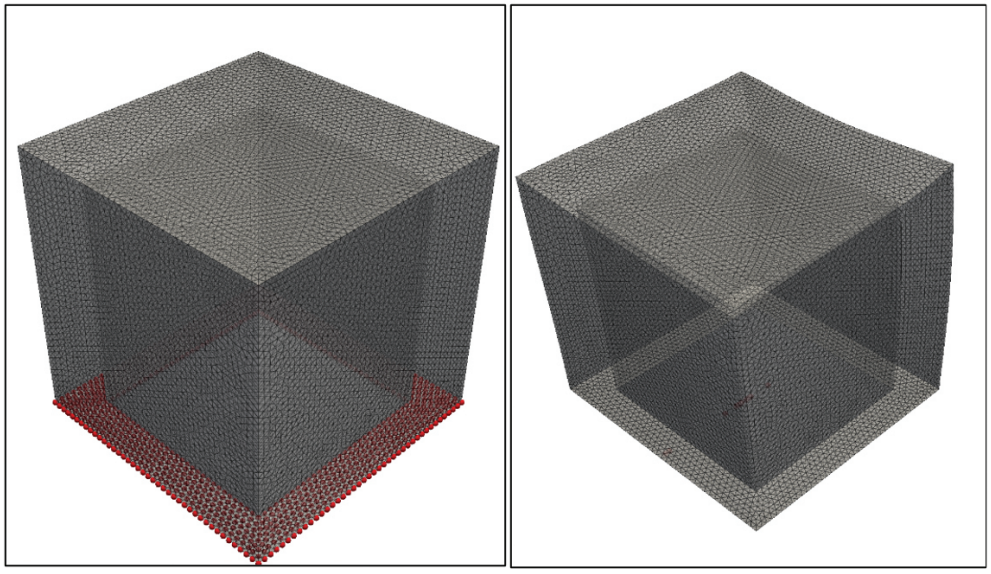


Fig. 7 Left: Resting (unloaded) configuration of the box. Red dots denote fixed nodes. Right: Deformed configuration of the box at time $t = 0$.

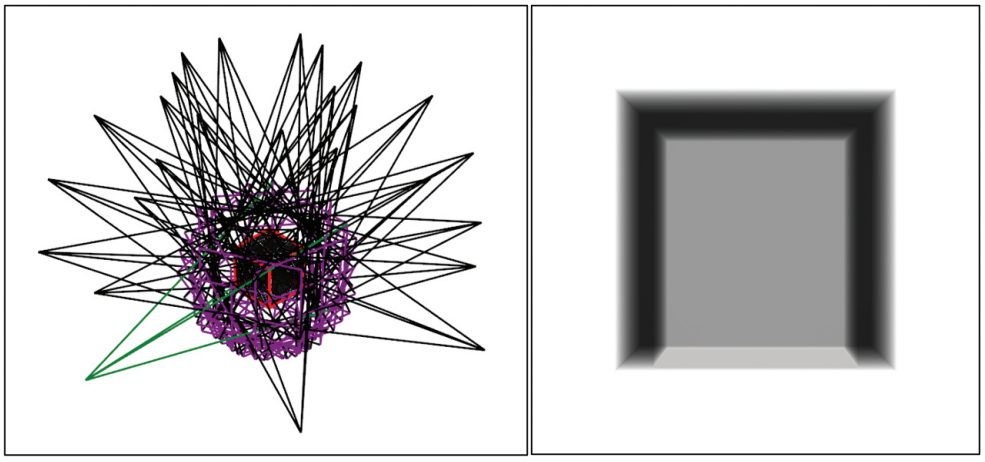


Fig. 8 Left: Imaging configuration. Pyramids represent imagers with x-ray sources at the tip and scintillator plates (purple) at the base. Red cube represents the voxel grid used to render images. Right: Radiograph rendered from x-ray imager picture in green in the left image.

Static Reconstruction

The static reconstruction procedure described in Figure 1 was applied to the simulated radiographs of the box in its rest configuration. Figure 9 shows two novel viewpoints of the reconstructed static attenuation grid, demonstrating that the geometry was recovered effectively.

Deflection Field Characterization / Dynamic Reconstruction

When attempting to learn the nodal displacement vector $\boldsymbol{u}(t)$ directly, we quickly noticed that the optimizer converges to high-spatial-frequency local minima that do not accurately reflect the behavior of the structure. As such, some degree of spatial smoothing was required in order to allow the optimization to converge to a set of nodal displacements that effectively characterized the motion of the structure. Since the dynamics of linear structures are typically sparse with respect to a modal basis, we assume that $\boldsymbol{u}\,(t) = \sum_{i=1}^{8} \boldsymbol{\phi}_i q_i(t)$ where $\boldsymbol{\phi}_i$ are the first 8 modes of the structure. For homogeneous structures, the mode shapes can be obtained with knowledge only of the structure's geometry. Note that these mode shapes are only used

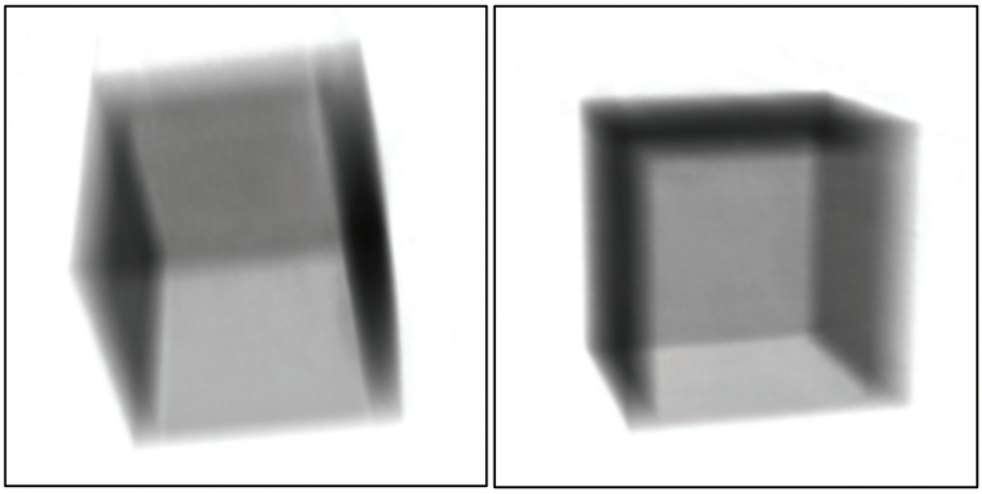


Fig. 9 Images rendered from the reconstructed attenuation field. Neither viewpoint was present in the data. Left: camera positioned beneath the box, looking up. Right: camera positioned beside the box, looking slightly down.

to smooth the learned displacement fields. In principle, for an inhomogeneous structure, the mode shapes corresponding to a materially homogeneous structure of the same shape could theoretically serve the same purpose, so long as enough of these modes are included to describe the observed dynamics of the structure.

Figure 10 below compares the true deflection at two timesteps to the deflection recovered using our method. Figure 11 performs a similar comparison using the modal coordinates of the structure.

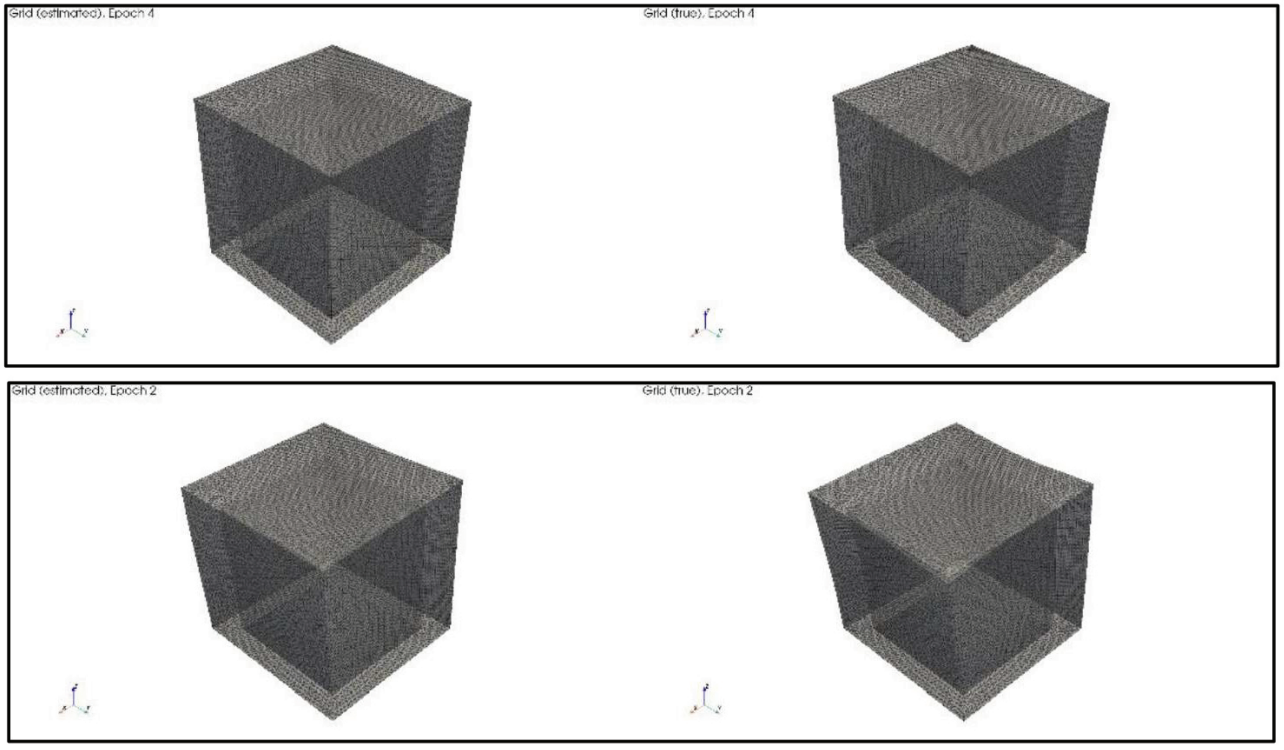


Fig. 10 True deflection (right) vs. recovered deflection (left) at timesteps 6 (top) and 17 (bottom).

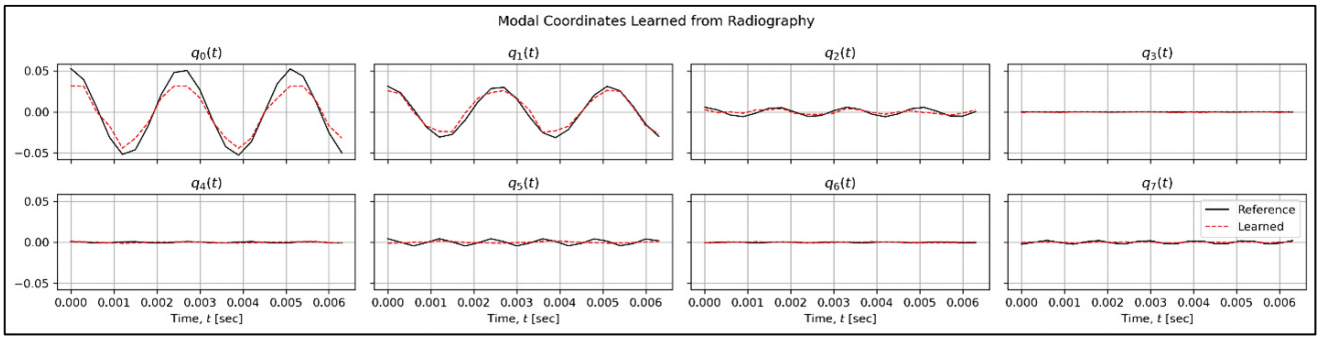


Fig. 11 True vs. recovered modal coordinates. Black line represents modal coordinates obtained by projecting the simulated nodal deflection history onto a modal basis. Red line represents modal coordinates recovered using our method.

System Identification

Having obtained the structure's modal coordinates at each time step, we apply the Rayleigh quotient method described in Section II-D to estimate the structure's Young's modulus (E), assuming that we know its density (ρ). We obtained $E_{\text{est}} = 67.16$ GPa. This is only a 4.05% difference with respect to the true value of $E_{\text{true}} = 70.00$ GPa.

Conclusion & Future Work

We have developed a framework for characterizing dynamic structural deformations from time-series radiography data. Additionally, we have demonstrated the framework's effectiveness on a synthetic experiment, assuming material linearity and homogeneity. While these initial results are promising, further work is still required to develop the framework to the point where it can be used on real experiments. Namely, steps must be taken to reduce the number of imager views necessary to characterize the displacement field, and to reduce the frequency at which these imagers must fire.

Due to time constraints, we did not experiment with different numbers of imagers and their placements. We expect that the current framework would likely work just as well with far fewer imager positions and intend to investigate this as future work. However, with fewer imagers present, it would likely become increasingly important to optimize their positioning such that they can observe the dynamics of the structure.

We believe that the second point may be addressed by introducing a Sinusoidal Representation Network (SIREN) to represent the modal coordinate vector. In the current procedure, the deflection vector associated with a given time step is only informed by images captured at that time step. By implicitly representing the deflection vector as a function of time using a SIREN, images captured at all time steps could influence the predicted nodal deflection vector at all other timesteps. This would introduce some complications in training the neural network, such as memory consumption and gradient variance (which we observed when we tried to implement this idea ourselves). However, if these issues could be effectively addressed, it would likely succeed in reducing the required firing rate of the x-ray sources by allowing for asynchronous imaging. For example, suppose there were 10 x-ray sources, each firing at 10 Hz. The sample overall sample rate could be achieved by firing each source at 1 Hz, with each source firing 0.1 seconds after the last. As an additional bonus, this method of representing nodal displacements would allow for less strict synchronization between imagers.

Finally, the underlying method of 3D representation proposed in this work shares many similarities with NeRF-style methods, specifically Direct Voxel Grid Optimization [9]. Our approach (with a few modifications) would likely work quite well with visible-light imagers for tasks that do not require knowledge of the behavior of internal structures.

References

1. S. W. Doebling, C. R. Farrar and M. B. Prime, "A summary review of vibration-based damage identification methods," *Shock and vibration digest*, pp. 91-105, 1998.
2. Y. Yang, C. Dorn, T. Mancini, Z. Talken, G. Kenyon, C. Farrar and D. Mascareñas, "Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification," *Mechanical Systems and Signal Processing*, pp. 567-590, 2017.
3. A. S. Poozesh, Z. Mao, C. Niezrecki and Peyman, "Vibration-based damage detection in wind turbine blades using Phase-based Motion Estimation and motion magnification," *Journal of Sound and Vibration*, pp. 300-318, 2018.

4. M. Southwick, Z. Mao and C. Niezrecki, “Volumetric Motion Magnification: Subtle Motion Extraction from 4D Data,” *Measurement*, p. 109211, 2021.
5. D. P. Rohe, J. Miers and Z. M. P. Avitabile, “Seeing Beyond the Surface – Structural Dynamics Identification through Multi-View Radiographic Imaging,” in *International Modal Analysis Conference*, 2025.
6. D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
7. F. P. Vidal and P. F. Villard, “Development and validation of real-time simulation of X-ray imaging with respiratory motion,” *Computerized Medical Imaging and Graphics*, pp. 1-15, 2016.
8. J. D. Frankel and S. Hoffman, “Numerical Methods for Engineers and Scientists,” in *Numerical Methods for Engineers and Scientists*, New York, CRC Press, 2001, pp. 141-145.
9. C. Sun, M. Sun and H.-T. Chen, “Direct Voxel Grid Optimization: Super-fast Convergence for Radiance,” CoRR, 2021.

Chapter 4
Mode-Shape Identification of Representative Samples with Digital Image Correlation

Devin Binns, Baiza Mand, Mason Rollans, Peter Beck, Christopher Teti, Kennan Hodovic, and Jim Declerk

Abstract Experimental modal analysis and environmental vibration tests are often used to characterize the modal parameters of a mechanical structure, namely the resonant frequencies, mode shapes, and damping ratios. Typically, these tests rely on attaching sensors to the structure to measure accelerations, which is then processed to extract the modal parameters. There are often testing limitations related to the number of sensors, placement, attachment method, and data acquisition channels needed when using accelerometers. Digital image correlation (DIC) has the potential to replace accelerometers in certain situations; however, there are still questions related to its validity in modal testing and correlation to modal testing performed with accelerometers. In this study, DIC was used to measure displacements in an object for the determination of modal parameters and then compare them to the modal parameters obtained from the traditional method of mounting accelerometers. A flat aluminum plate was tested via impact excitation and results from measurements with DIC were compared to a finite element solution and testing with accelerometer measurements. The measured mode shapes, which qualitatively agreed with both the accelerometer results and the theoretical model, provided higher spatial fidelity than accelerometers. Results produced with DIC measurements yielded natural frequencies with a similar accuracy as accelerometers. Tests with and without accelerometers showed a consistent 3% difference in natural frequencies due to accelerometer mass loading as a source of error.

Keywords Mode shape identification · Experimental modal analysis · Digital Image Correlation (DIC) · Impact testing · Finite Element · Image Processing

List of Notations and Abbreviations: Digital Image Correlation (DIC), Laser Doppler Vibrometry (LDV), Finite Element Analysis (FEA), Data Acquisition Unit (DAQ), Direct Current (DC), Integrated Circuit Piezoelectric (ICP), Transistor-Transistor Logic (TTL), Comma Separated Variable (CSV), Frequency Response Function (FRF), Fast Fourier Transform

Devin Binns
Department of Aerospace and Mechanical Engineering, New Mexico State University, Las Cruces, New Mexico 88003
e-mail: dmbinns@lanl.gov

Baiza Mand
Department of Engineering Physics and Department of Computer Science, Colorado School of Mines, Golden, Colorado 80401
e-mail: mand@lanl.gov

Mason Rollans
Mechanical Engineering Department, Brigham Young University, Provo, Utah 84602
e-mail: mjrollans@lanl.gov

Peter Beck · Christopher Teti
Verification and Validation Division, Los Alamos National Laboratory, Los Alamos, NM 87545
e-mail: pmbeck@lanl.gov

Kennan Hodovic
Engineering Division, Los Alamos National Laboratory, Los Alamos, NM 87545
e-mail: kennan@lanl.gov

Jim Declerk
Department of Mechanical and Aerospace Engineering, Michigan Technological University, Houghton, Michigan 49931
e-mail: jdeclerck@mtu.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

(FFT), Auto Power Spectrum Density (APSD), Effective Independence Method (EFI), Frames Per Second (FPS), Signal to Noise Ratio (SNR), Modal Assurance Criterion (MAC)

Introduction

Structures in a service environment experience many vibrations and shocks which lead to wear, causing premature failure. As a result, many industries use vibrational testing to qualify and validate components that will be used in a service environment. Commonly, vibrational tests use accelerometers to record and collect a test article's dynamic response. These sensors can distort results due to the additional mass of the accelerometers on the test article. To avoid this issue, measurement devices that do not physically contact the test article must be used. Additionally, accelerometers provide sparse information about mode shapes since they only acquire data at a few discrete points. One alternative method to measure vibrational responses of an object is Stereo Digital Image Correlation (DIC). Stereo DIC utilizes two high-speed cameras oriented towards the test article at specified angles to track the location of a group of speckle marks located on the test article [1]. Modal measurements with DIC can enable testing in high-shock environments; with hazardous materials such as Beryllium, or high explosives; or in testing large components which would otherwise require an excessive number of accelerometers to characterize [2]. DIC can even be used on small components where mass loading becomes a more prominent issue [3].

Previous comparisons between DIC and Laser Doppler Vibrometry (LDV) [4, 5] have shown that both technologies can produce similar results to those from traditional accelerometer measurements. It was found that LDV yields more precise out-of-plane results and equally precise in-plane results when compared to DIC. However, LDV lacks the strain measurements that DIC provides, which is useful for qualification testing. DIC can provide instantaneous full-field measurements in contrast to LDV and accelerometer measurements. Scanning LDV can perform full-field measurements, however, the process of scanning does not allow for time-dependent testing, such as shock testing with an impact hammer. Warren et al. having conducted impact tests, stated that DIC provides a distinct advantage over LDV in the measurement of rigid body motion and large vibrational displacements corresponding to lower frequencies [6]. Further investigations into the validity of DIC for modal analysis have pointed out that high frequency vibrations, which are easily detectable by accelerometers, may not displace above the noise floor for DIC, limiting the maximum frequency measurable [4]. Javh et al. introduced a hybrid DIC-accelerometer approach where an accelerometer is used to find the resonant frequencies and damping ratios, enabling DIC to find the high-frequency mode shapes utilizing full-field data after combining the Least-Squares Complex Frequency method and the Least-Squares Frequency-Domain method [7]. Most of the previous studies discussed here studied DIC for use in forced-normal-mode testing. There exists minimal research on using DIC for impact tests.

Through this study, our aim was to further investigate the viability of DIC measurements through a comparison to accelerometer measurements and preliminary Finite Element Analysis (FEA) models on a simple 6061-T6 aluminum plate. We performed free vibrational tests using a modal hammer to excite the plate and measured the plate's response as it rang down using both a calibrated DIC system and accelerometers. Due to the nature of impact testing, we simultaneously activated and analyzed multiple modes. This differs from previous studies where only one mode shape was analyzed at a time using forced-normal-mode-excitations. We used natural frequencies, mode shapes, and damping ratios to compare all measurement methods and the FE model.

Methods

Test Apparatus

The test article was an aluminum 6061-T6 plate with dimensions of 7 in by 8 in by 3/16 in, seen in Figure 1. The flat rectangular shape was chosen to simplify the comparisons between measurement methods and to avoid symmetric modes that would exist with a square plate. We opted to investigate the out-of-plane dynamics of the plate. The two top corners of the plate were connected with fishing line to a rigid mounting fixture, while the two bottom corners were loosely attached to the fixture with bungee cords to remove excess rigid body motion and keep the plate within the cameras' focus and field of view. In the finite element model, this mounting condition was approximated with free-free boundary conditions.

Finite Element Modeling of Plate Modes

We developed a model of the aluminum 6061-T6 plate within Abaqus to give a preliminary estimate of the mode shapes and natural frequencies of the plate. We performed a mesh convergence study, and the model was found to converge well using quadratic shell elements with seven integration points and an element size of 2.5 mm.

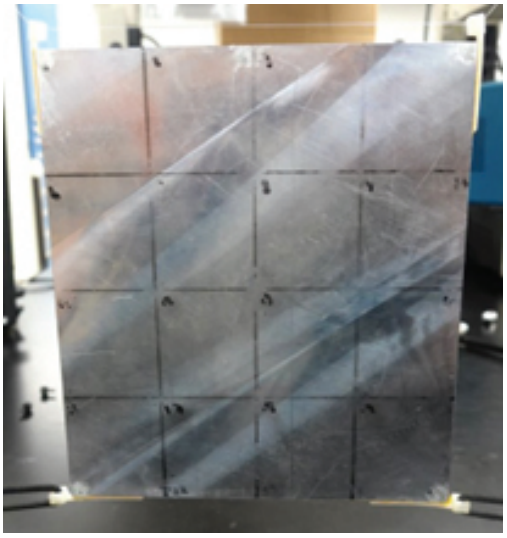


Fig. 1 Aluminum 6061-T6 rectangular test article

The first nine modes, shown in Figure 2, were used to compare DIC measurements and accelerometer measurements. Table 1 displays the natural frequencies associated with the first nine modes found by the FE model. These nine modes are all distinct, which allowed for clear confirmation on which modes we could identify within our experimental results.

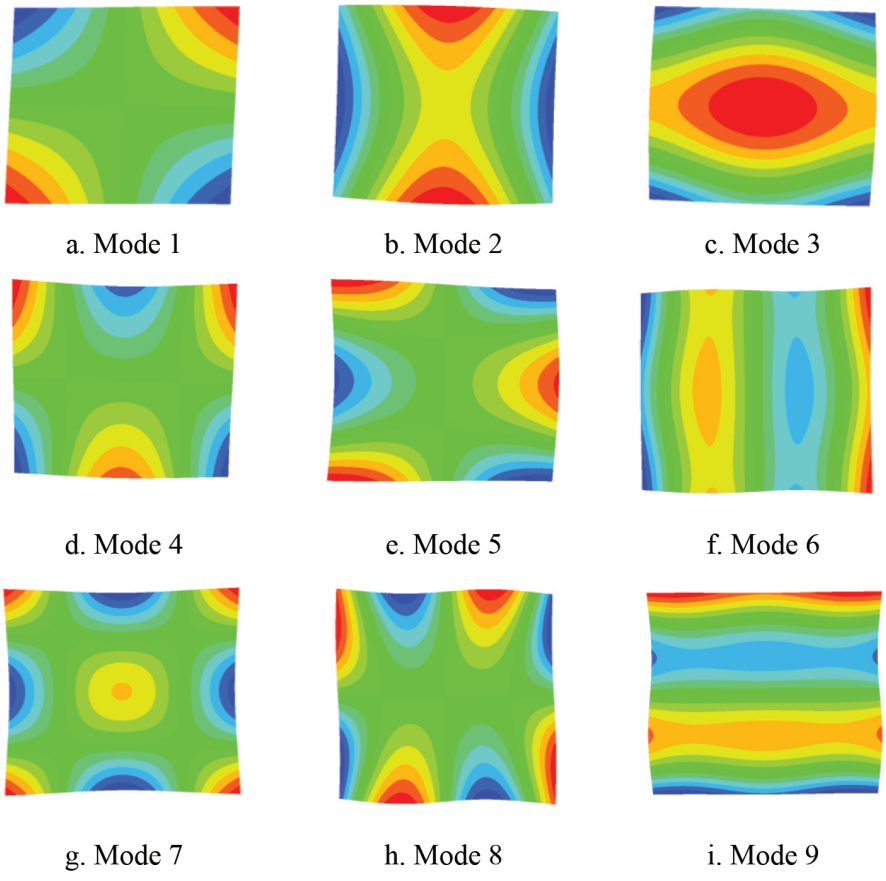


Fig. 2 Finite element model mode shapes

Table 1 Predicted natural frequencies from the finite element model.

Mode	1	2	3	4	5	6	7	8	9
Natural Frequency (Hz)	420	582	833	1040	1150	1710	1990	2070	2220

Accelerometer Setup

Using results from the FE model, we determined the optimal locations to mount five accelerometers. Locations were chosen to maximize the measured information about the first five mode shapes and therefore maximize our ability to distinguish between mode shapes from the accelerometer tests. To find the optimal accelerometer locations, we assigned a grid of 25 evenly spaced points on our Abaqus model as candidate locations and used the Effective Independence (EFI) method [8] to determine ideal accelerometer locations. EFI selected the four corners and the midpoint on the top edge of the plate, which we used as the accelerometer locations in our setup as shown in Figure 3.

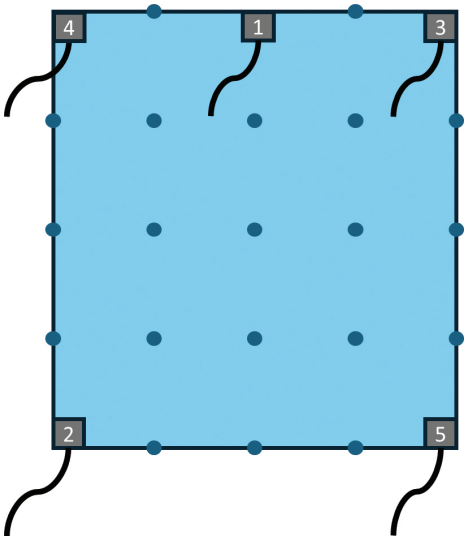


Fig. 3 Graphic of test article showing the grid of 25 candidate points and the optimized accelerometer placement based on EFI results

After mounting the accelerometers on the back side of the plate using modal wax, we performed a traditional roving hammer impact test by impacting the back of the plate at 25 points shown in Figure 3. For the 5 locations where accelerometers were attached, we impacted the front side of the plate to avoid directly impacting the accelerometers. A PCB Piezotronics SN22337 modal hammer was used to both trigger the DAQ and measure the impact force. Five PCB Piezotronics 356A01 triaxial accelerometers measured the output response after every hit. To sample from our sensors, we used a Siemens 16-channel LMS SCADAS Mobile DAQ and collected the sensor data with Simcenter Testlab.

DIC Setup

The experimental setup for DIC measurements is shown in Figure 4. To enable DIC measurements, we generated a random black and white speckle pattern in MATLAB and printed it on sticker paper. We applied the sticker paper to the face of the plate which faced the cameras. We mounted two color Photron Fastcam SA-Z high speed cameras to an optical breadboard and angled them at approximately 30° from each other. We fitted both cameras with a Canon 50mm lens set to an aperture of f/7.1, allowing an effective depth of field of approximately 42 mm for each camera, and ensuring that the plate would remain in focus during testing. We synchronized the cameras to each other by triggering them both via a single trigger input channel. To illuminate the test article during testing, we used two Veritas Constellation 120E lights with a maximum output of 12,700 lumens.

During testing, the impact hammer triggered both the cameras and the DAQ. Typically, the hammer operates at a constant 10 VDC, which briefly spikes during impact. To enable triggering of the measurement devices, we designed a custom circuit to remove the DC offset voltage and ensure the spike is large enough to trigger the devices, but not large enough to damage them. This circuit consisted of a buffer to avoid altering the measured hammer voltage, followed by a high pass filter to

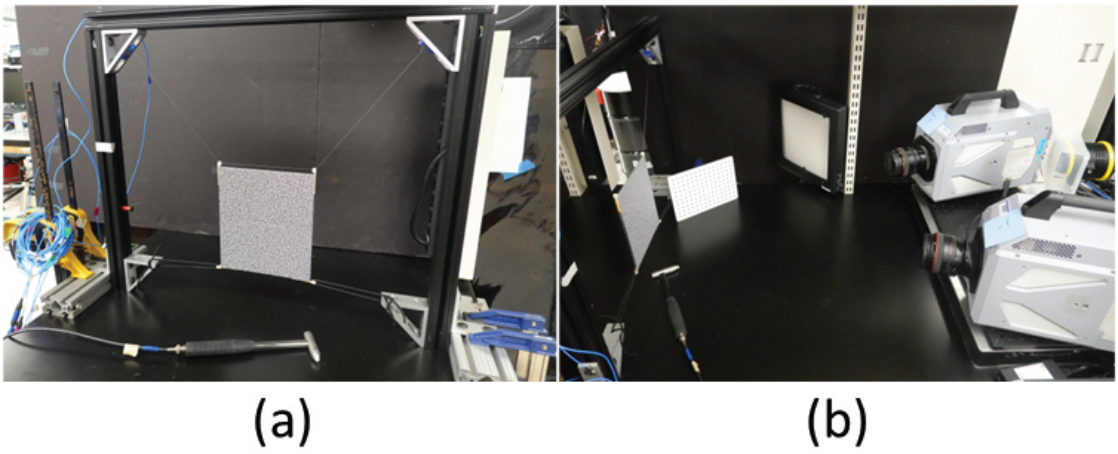


Fig. 4 Experimental test setup, including (a) the mounted test article and modal hammer used for excitation, and (b) cameras and lights pointed at the test article

remove the DC bias of the ICP voltage from the DAQ. Finally, the circuit amplified the signal to ensure the output voltage would exceed the 2.5 V trigger voltage recognized by the cameras and the DAQ. For the DIC recording, we set a pre-trigger of approximately 100 frames to capture the start of the impulse. To synchronize the timing of each measurement between hits, we applied a similar pre-trigger of the same length in Testlab.

We developed a workflow for conducting impact tests during a DIC recording to facilitate a comparison between DIC and accelerometer measurements. This method leveraged the use of the filter circuit so that successive hammer impacts would trigger separate DIC recordings which were partitioned within the camera's internal storage, allowing us to execute and record multiple impacts along the plate's surface. We triggered the DAQ using the same hammer impacts to simultaneously record the force data from the modal hammer, ensuring consistent and synced data recordings which could later be used in the development of Frequency Response Functions (FRFs). One limitation of this method was recording time. All impact events were stored on the camera's internal 64 GB memory, allowing for a total of 65,536 frames for the entire roving hammer test, assuming 8-bit color and 1 mega-pixel frames. It is possible to capture more frames by downloading images between hammer strikes; however, due to time constraints, we limited one full roving hammer test to the memory capacity of the cameras. Due to these recording limitations, we selected only 9 impact locations by excluding all edge impact locations. This increased the number of frames captured per hammer strike and minimized the probability of double-impacts which were more likely to occur on edge-impacts. An additional benefit to the reduced number of hammer strikes is that all strikes could be performed on the same side of the plate as the accelerometers, enabling testing with the accelerometers attached.

We recorded nominal DIC measurements at 5000 frames per second (FPS), which allowed for a total recording time of 13.1 seconds, or 1.46 seconds per impact location. To quantify mass loading caused by the attached accelerometers, we conducted two separate DIC tests: one test with attached dormant accelerometers, referred to as "DIC with mass loading", and one test without attached accelerometers, referred to as "DIC without mass loading". Additionally, to investigate the limits of frequency measurements and displacement resolution for our current setup, we conducted a third DIC test, referred to as "DIC at 10k FPS", with no accelerometers attached, at twice the frame rate of the previous tests, 10,000 FPS. For this test, we recorded only 4 impact locations, with 1.6 seconds per impact location. Because the image exposure time decreases as the frame rate increases, we were unable to increase the frame rate above 10,000 FPS with our current lighting setup. More lights, and/or brighter lights would enable higher frame rate recordings.

The processing steps are shown in Figure 5. First, we grayscaled the images and renamed them to be compatible with Vic-3D. Next, we processed the images in Vic-3D using the settings shown in Table 2. This yielded a series of Comma Separated Variable (CSV) files with information about the resonant response of the plate. We combined these files into one large CSV file containing only the out-of-plane displacement of each point at each time step. From this, we applied an exponential window to the displacement data, took the fast Fourier transform (FFT) at each point on the plate, and multiplied by $-\omega^2$ to convert to acceleration. We also applied the same window to and took the FFT of the hammer force data. With these frequency spectra, we used the H1 estimator method to calculate FRFs. Plotting the FRFs allowed us to locate the resonant frequencies of our plate. Additionally, we visualized the plate's mode shapes by plotting the modal vectors, the imaginary components of the FRFs at the resonant frequencies, on a 2D heat map.

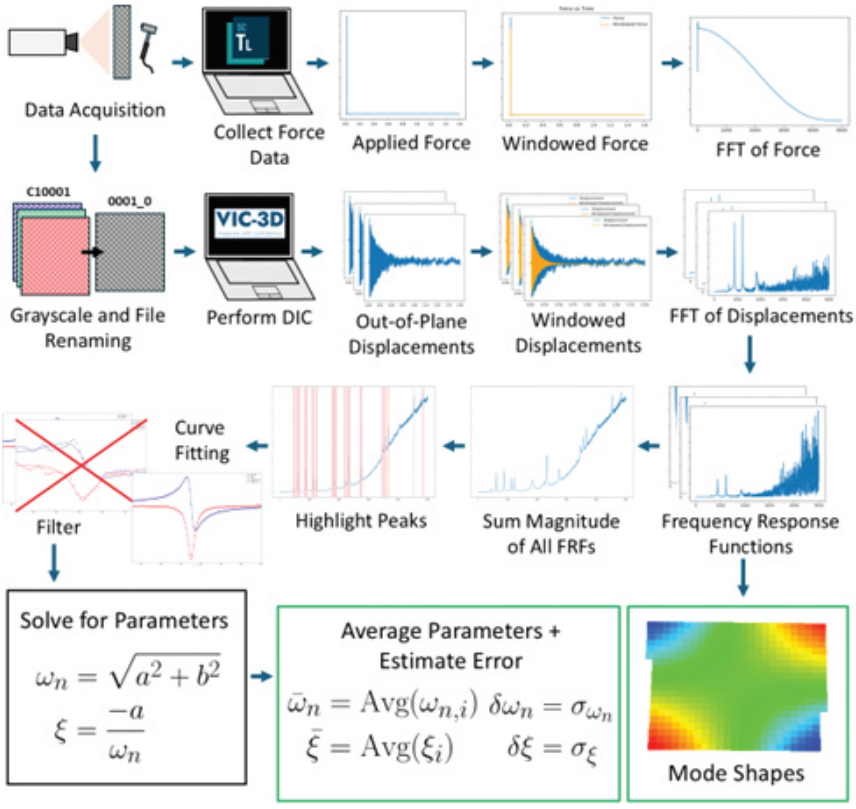


Fig. 5 DIC modal analysis workflow including data acquisition and the processing of both the camera and hammer data to extract modal information

Next, we plotted the sum of the magnitudes of the FRFs from each point on the plate and from each impact trial. This plot enabled us to highlight a region-of-interest around each peak so we could locally curve-fit to each peak in each FRF separately. We individually curve-fit the real and imaginary component of each peak in each FRF to the real and imaginary component of Equation 2.4.1, where A, B, a, b, $y_{0,Real}$, and $y_{0,Imag}$ are curve-fit parameters, ω is the independent variable frequency, i is the imaginary unit, and H is the transfer function. Each fit provided parameters which we used to calculate the damping ratio, ξ, and the natural frequency, ω_n, of the peak at that point for that trial, according to Equations 2.4.2 and 2.4.3 [9]. Poor fits were often caused by an FRF having a low signal-to-noise ratio for that peak, meaning the FRF contained little or no useful information about the modal parameters for that mode; therefore, we filtered out any results where the parameters a and b had an average standard error of greater than 5%. To estimate the damping ratio and natural frequency of a mode, we averaged together all the remaining results for that mode. We estimated our uncertainty in the damping ratio and natural frequency of each mode as the standard deviation of their corresponding distributions.

Table 2 Vic-3D settings used in this study.

Software setting	**Value**
Correlated solutions	Vic-3D v 9
Subset	39
Step size	30
Pixel size	0.2 mm/pixel
Minimization	ZNSSD
Interpolant	8-Tap
Image pre-filtering	None
Speckle size	1.25 mm

$$H = \frac{A + iB}{i(\omega - b) - a} + \frac{A - iB}{i(\omega + b) - a} + y_{0,\ Real} + iy_{0,\ Imag} \tag{1}$$

$$\omega_n = \sqrt{a^2 + b^2} \tag{2}$$

$$\xi = \frac{-a}{\omega_n} \tag{3}$$

Results and Discussion

Natural Frequency Results

To compare between measurement methods, one of the parameters we chose to investigate was our test article's natural frequencies. For FEA, we found the natural frequencies by conducting a frequency step analysis on our meshed model. For accelerometer testing, we used the roving impact hammer test described in Section 2.3 to collect data. Testlab produced FRFs and we conducted the post-processing steps described in Section 2.4 to find natural frequencies from a collection of FRFs. We used the steps described in Section 2.4, as opposed to Testlab's built-in post-processing, to ensure accelerometer and DIC data would be analyzed consistently. For DIC testing, we used the roving impact hammer workflow and post-processing steps described in Section 2.4 to find natural frequencies.

Figure 6 shows the difference between data from DIC and accelerometers for the first nine natural frequencies. The uncertainty from both accelerometer measurements and DIC measurements is propagated into the plotted percent differences.

For the DIC with mass loading test, there is very little discrepancy between measured natural frequency values for most modes, with absolute percent differences falling within a range of 0% to 1.635% and an average absolute percent

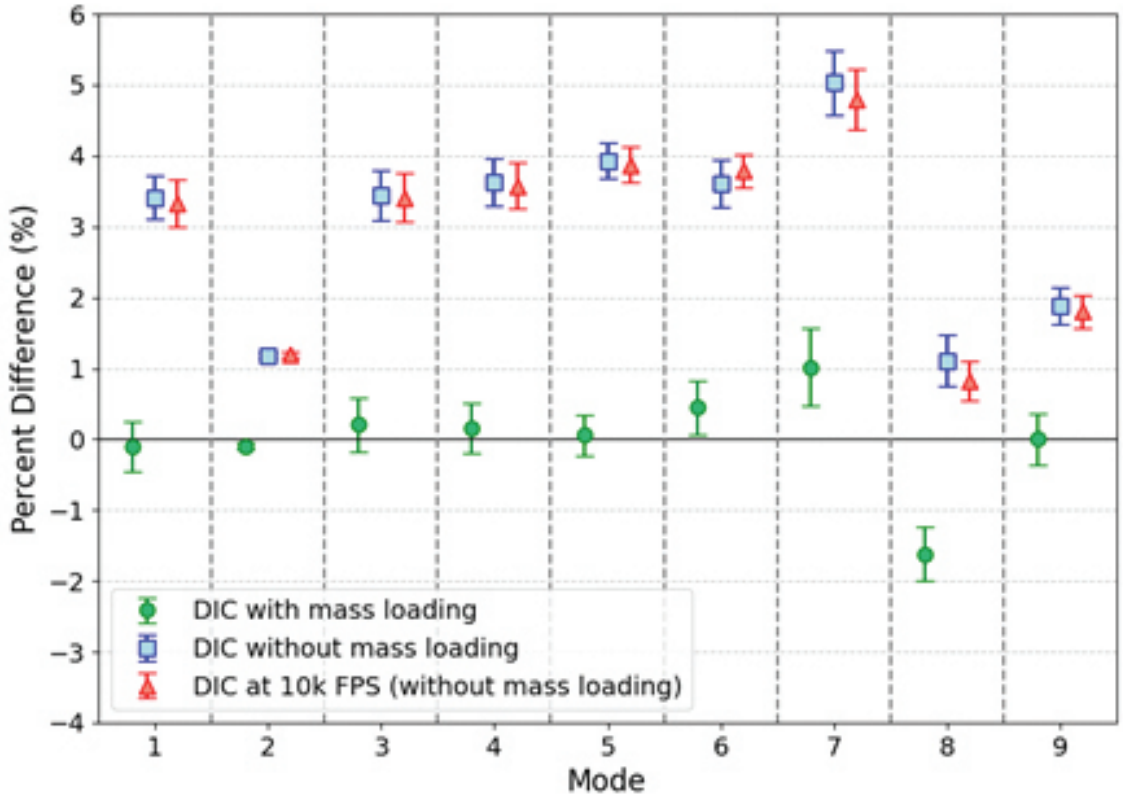


Fig. 6 Plot of percent differences of natural frequencies found using each DIC test from values found using the accelerometer test

difference of 0.414%. An outlier in this dataset was mode 8, which had a low signal-to-noise ratio (SNR) for the DIC with mass loading test. This caused zero curve fits to meet the filtering criteria during post-processing, so we chose to calculate modal parameters using only the best 1% of curve-fits as determined by the average standard error of their fit-parameters a and b. Aside from mode 8, there is a strong agreement between the accelerometer and DIC measurements of the natural frequencies.

On average, tests measured with DIC, with five accelerometers concurrently mounted, resulted in an average 3.017% lower natural frequencies than tests performed with DIC and no accelerometers mounted. We believe this offset is a result of mass loading caused by accelerometers. Tension and friction in the accelerometer cables could also contribute to the difference. There is strong agreement between the DIC test at 10K FPS and the DIC without mass loading, further supporting the hypothesis that mounted accelerometers alter the plate's natural frequency.

Table 3 Natural frequency measurements from FEA, accelerometers, and DIC.

	FEA	Accelerometer	DIC with mass loading	DIC without mass loading	DIC at 10k FPS
Mode	ω_n (Hz)	ω_n (Hz)	ω_n (Hz)	ω_n (Hz)	ω_n (Hz)
1	420	443 ±1	442.6 ±0.8	458.4 ±0.4	458.0 ±0.7
2	582	607.8 ±0.2	607.2 ±0.2	614.9 ±0.1	615.0 ±0.1
3	833	849 ±3	851 ±1	879.1 ±0.5	878.9 ±0.4
4	1040	1090 ±4	1092 ±1	1130.2 ±0.5	1129.7 ±0.3
5	1150	1184 ±3	1185 ±2	1231.8 ±0.4	1231.0 ±0.2
6	1710	1773 ±2	1780 ±6	1838 ±6	1841 ±4
7	1990	2082 ±8	2103 ±8	2189 ±5	*2184 ±4
8	2070	2220 ±6	*2184 ±6	2244 ±6	2238 ±2
9	2220	2288 ±5	2288 ±6	2332 ±3	2330 ±1
10	2580	2616 ±--	---	---	2748 ±2
11	3138	3340 ±--	---	---	*3474 ±1
12	3317	3497 ±--	---	---	*3559 ±8
13	3398	3634 ±--	---	---	*3670. ±4
14	3752	3859 ±--	---	---	*4061 ±7
15	4326	4447 ±--	---	---	*4543 ±6

* Curve-fitting filter criteria altered - Result may be unreliable.

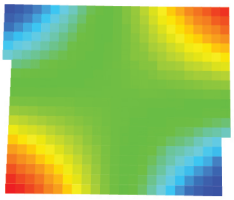
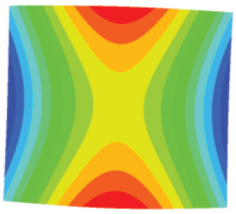
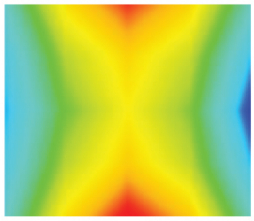
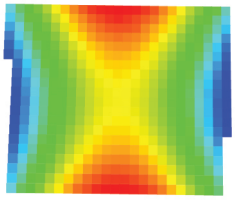
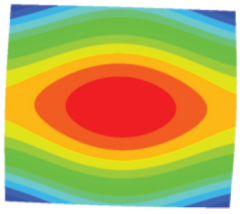
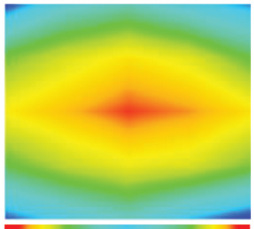
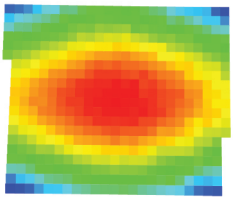
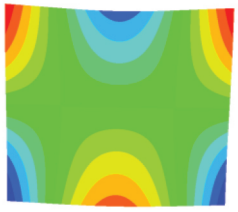
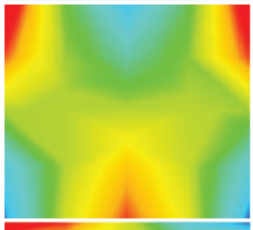
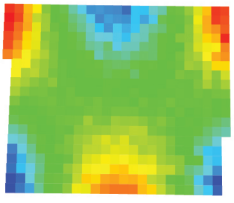
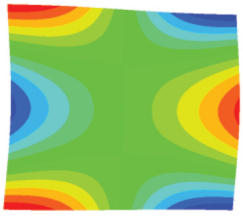
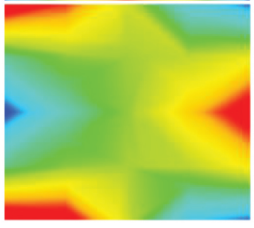
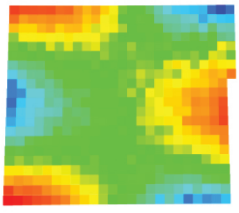
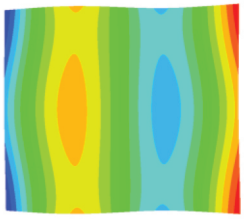
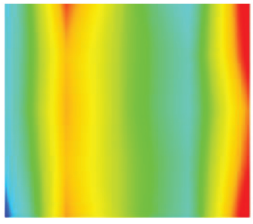
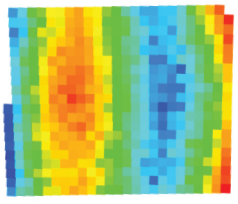
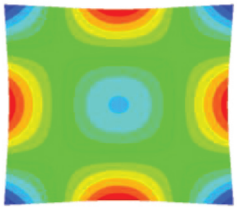
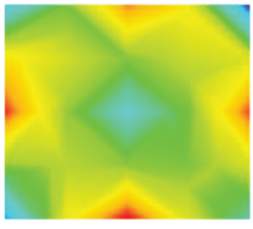
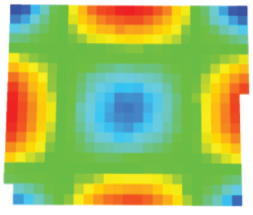
Mode Shape Results

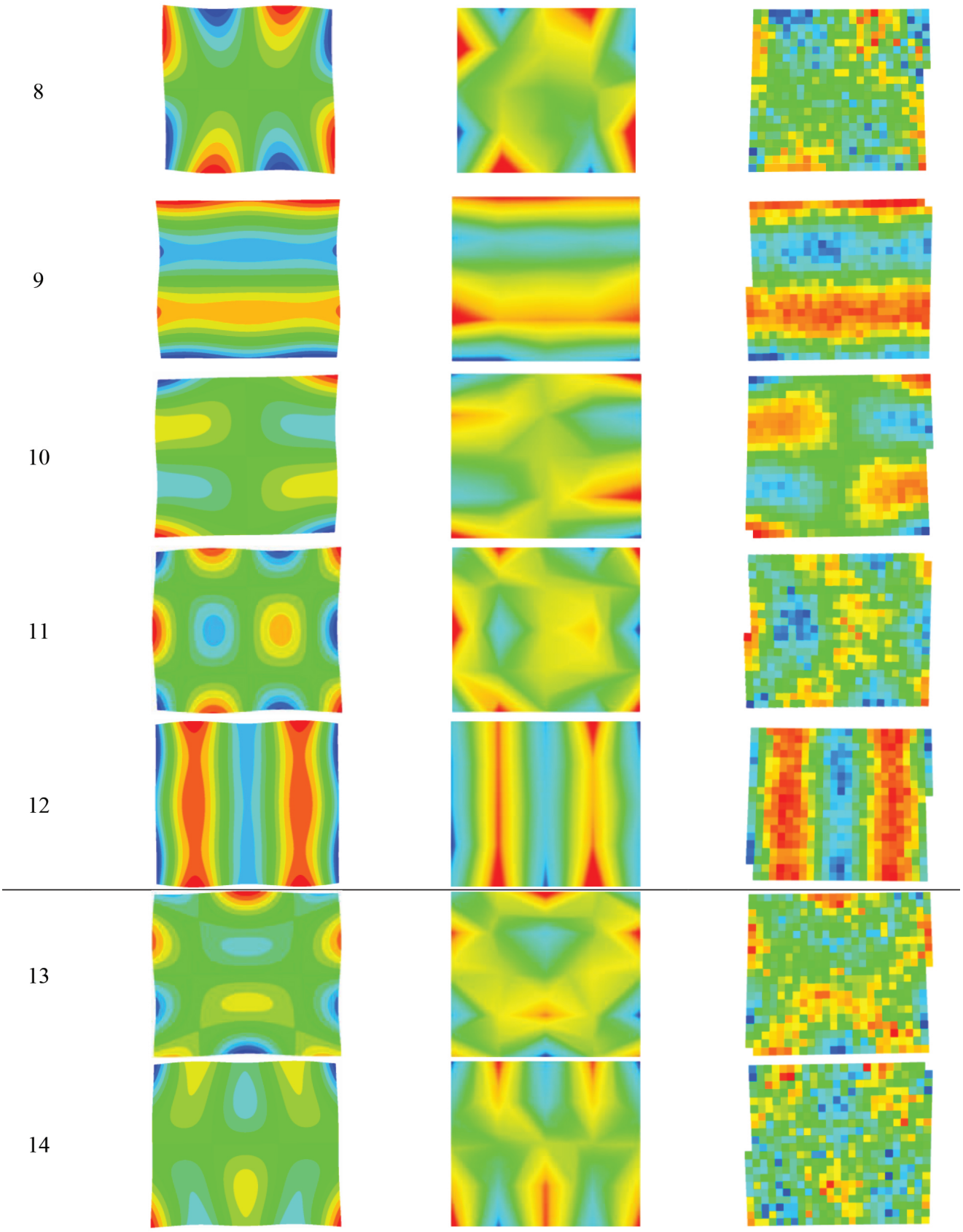
Mode shapes from FEA, accelerometer measurements, and DIC measurements were obtained and are visually compared. Accelerometer results were obtained using Testlab. Because we only sampled the plate at 25 hammer impact points, Testlab interpolated the amplitude between nodes to plot smoother mode shapes. DIC results were obtained using the steps described in Section 2.4. DIC's full field measurement yields mode shapes with higher resolution, so no interpolation was necessary to visualize the mode shapes. It is important to note that for accelerometer results, contributions from all impact locations on the plate contribute to the mode shape plots, while for DIC results, plots are taken from the highest amplitude peak across all impact events from a roving impact hammer test. Further work may be done to incorporate contributions from all impact locations with DIC testing, as well. Mode shape plots are qualitatively compared between various methods in Table 4.

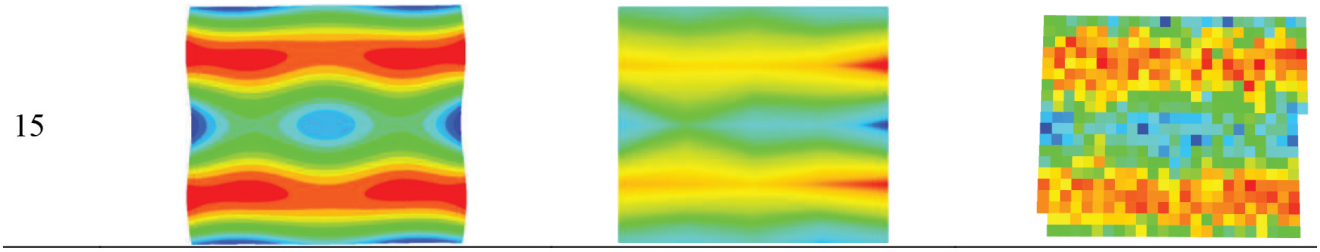
From the various mode shapes shown in Table 4 we observe agreement between most modes from all measurement methods. FEA and accelerometer results show strong agreement throughout, while the noise of the DIC mode shapes increases as the frequency increases. This correlation is due to the higher frequency modes having a lower magnitude of displacement; the displacement magnitudes are close to the noise floor of DIC measurements. This effect is especially apparent in mode 8 because our DIC roving hammer test didn't impact any antinodes. When measuring with DIC the test article can be impacted with significantly more force than with accelerometers, due to the risk of damaging or overloading the sensors. This aided in resolving the higher mode shapes by increasing their deflection above the noise floor of DIC.

Overall, it is posited that if smoothing algorithms like ones used within Testlab were used, DIC results could obtain a higher qualitative agreement between all modes investigated. Future analysis would include the creation of cross-MAC matrices comparing all three measurement methods to get more objective and conclusive results.

Table 4 Mode shape plots from FEA, accelerometers, and DIC.

Mode	FEA	Accelerometers	DIC
1			
2			
3			
4			
5			
6			
7			


8
9
10
11
12
13
14



Damping Ratio Results

The final parameter we chose to compare was our test article's damping ratios. The methods for finding natural frequencies, described in Section 3.1, were also used for finding damping ratios from both accelerometer and DIC measurements.

To compare damping ratios, we created Figure 7 which shows the percent difference of the first nine damping ratios found using each DIC test from values found using the accelerometer test. We propagated the uncertainty from accelerometer measurements and DIC measurements into the plotted percent differences.

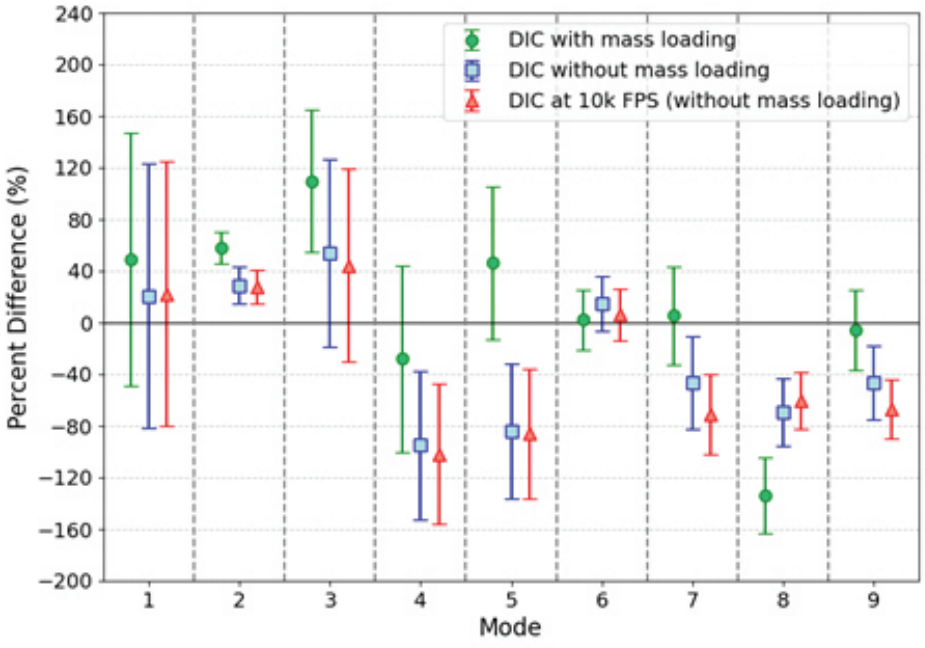


Fig. 7 Plot of percent differences of damping ratios found using each DIC test from values found using the accelerometer test

There was high variability when obtaining the values for damping ratios from different FRFs, resulting in large error bars as shown in Figure 7. The DIC with mass loading results tended to vary, either being less than or greater than the accelerometer baseline value. Due to this variability and the limited number of tests conducted, it is difficult to draw any specific conclusions comparing measurements conducted by accelerometers and DIC.

Our first hypothesis to explain the discrepancy between the accelerometer and DIC with mass loading measurements is that the addition of accelerometers can cause variability in the measured damping ratios. This is believed to not be caused by the mass loading itself but by the method by which the accelerometers are attached and the cable management associated with their attachment. PCB Piezotronics Model 080A109 petro wax was used as the mounting adhesive for experiments that required accelerometers. The accelerometer cables were secured with masking tape to prevent cable slap and tension during testing. However, it is still believed that the accelerometers may have introduced variable damping into the system. Further experiments would need to be conducted to either confirm or deny this hypothesis. Alternatively, the accelerometer measurements of damping ratios average results from all 25 impact locations, but the DIC measurements average results from hundreds of measurement locations. Since damping ratios may not be the same at all locations on a non-ideal test

article, calculating results using different measurement locations would most likely yield different results. For the two DIC tests conducted without the presence of accelerometers, we see strong agreement between the results, demonstrating the repeatability of this measurement method. It shows that DIC measurements conducted during impact tests can be used to precisely measure damping ratios.

Table 5 Damping ratio measurements from accelerometers and DIC.

	Accelerometers	DIC with mass loading	DIC without mass loading	DIC at 10k FPS
Mode	ξ	ξ	ξ	ξ
1	0.010 ±0.008	0.013 ±0.002	0.010 ±0.001	0.010 ±0.002
2	0.0026 ±0.0003	0.0047 ±0.0003	0.0035 ±0.0003	0.0034 ±0.0002
3	0.010 ±0.002	0.011 ±0.001	0.0054 ±0.0005	0.0049 ±0.0005
4	0.008 ±0.007	0.007 ±0.001	0.0034 ±0.0005	0.0031 ±0.0003
5	0.009 ±0.004	0.010 ±0.001	0.0024 ±0.0005	0.0024 ±0.0002
6	0.007 ±0.001	0.012 ±0.002	0.014 ±0.002	0.013 ±0.002
7	0.017 ±0.005	0.016 ±0.003	0.009 ±0.002	*0.007 ±0.001
8	0.008 ±0.002	*0.002 ±0.001	0.005 ±0.001	0.0054 ±0.0009
9	0.009 ±0.002	0.007 ±0.001	0.0044 ±0.0009	0.0035 ±0.0004
10	0.0116 ±--	---	---	0.0039 ±0.0007
11	0.0069 ±--	---	---	*0.0036 ±0.0006
12	0.0023 ±--	---	---	*0.0018 ±0.0009
13	0.009 ±--	---	---	*0.0015 ±0.0008
14	0.0089 ±--	---	---	*0.0013 ±0.0007
15	0.0129 ±--	---	---	*0.0014 ±0.0004

* Curve-fitting filter criteria altered - Result may be unreliable.

Quantification of Mass Loading Effects

Recognizing that there could be differences between values as an effect of accelerometer inclusion or exclusion, we broke our tests into two distinct categories: with and without mass loading due to accelerometers being mounted. It was found that there was a distinct difference between the results of DIC tests that did and did not include attached accelerometers. This shows that there was a significant effect of mass loading caused by the accelerometers. This was most apparent in the natural frequencies where values tended to be relatively consistent between tests except when comparing mass loading scenarios. Here, as stated before, mass loading induced by accelerometers caused, on average, a decrease of 3.017% in the plate's natural frequencies. From damping ratios, it was also seen that there was a change caused by accelerometers. However, it is difficult to conclude the exact nature of the change, due to the large variability of results from tests with accelerometers attached. Further testing would likely lead to stronger conclusions.

Looking qualitatively at Figure 8, which shows an overlay of two FRFs between two DIC tests, one with and one without mass loading, we can see an apparent shift in natural frequencies and change in damping in the peaks. More specifically, the DIC with mass loading FRF shows a slightly lower resonant frequency for all the modes presented, which is consistent with the test article having a larger mass. The peaks of the FRF from the DIC without mass loading test tend to be more prominent with narrower widths indicating less damping. The increased damping that occurs in the DIC with mass loading tests comes from the added accelerometers along with the wax used to adhere them and the cables.

Tradeoffs Between DIC and Accelerometers

One of the benefits of DIC is its high spatial resolution. DIC can measure hundreds of points simultaneously which would require many accelerometers and DAQ channels to match. Additionally, using a high quantity of accelerometers can introduce even more significant mass loading and therefore can often be infeasible. This shortcoming of accelerometers is commonly addressed by attaching a small number of sensors and exciting the test article at many points and leveraging the reciprocity of linear systems to characterize the article's dynamics with higher fidelity. While this method requires few accelerometers,

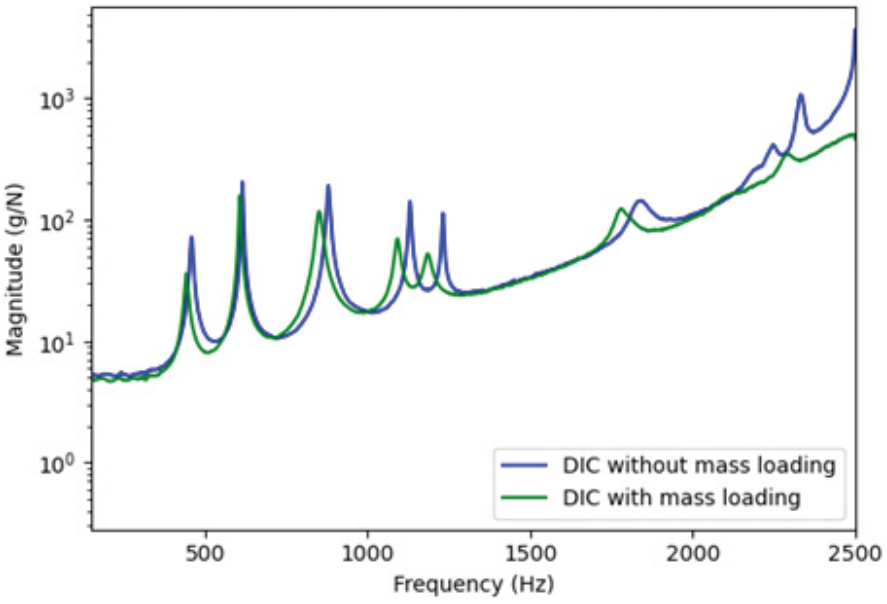


Fig. 8 Frequency response function plot from both DIC with and without mass loading

Table 6 Comparison of natural frequency and damping ratio measurements between DIC with and without mass loading.

	DIC with mass loading		DIC without mass loading		Percent difference	
Mode	ω_n (Hz)	ξ	ω_n (Hz)	ξ	ω_n %	ξ %
1	443.0 ±0.8	0.013 ±0.002	458. ±0.4	0.010 ±0.001	3.417	-26.09
2	607.8 ±0.2	0.0047 ±0.0003	614.9 ±0.1	0.0035 ±0.0003	1.161	-29.27
3	849 ±1	0.011 ±0.001	879.1 ±0.5	0.0054 ±0.0005	3.484	-68.29
4	1090 ±1	0.007 ±0.001	1130.2 ±0.5	0.0034 ±0.0005	3.621	-69.23
5	1184 ±2	0.010 ±0.001	1231.8 ±0.4	0.0024 ±0.0005	3.957	-122.6
6	1773 ±6	0.012 ±0.002	1838 ±6	0.014 ±0.002	3.600	15.38
7	2082 ±8	0.016 ±0.003	2189 ±5	0.009 ±0.002	5.010	-56.00
8	*2220 ±6	*0.002 ±0.001	2244 ±6	0.005 ±0.001	*1.075	*85.71
9	2288 ±6	0.007 ±0.001	2332 ±3	0.0044 ±0.0009	1.905	-45.61

* Curve-fitting filter criteria altered - Result may be unreliable.

it is more time-consuming. On the other hand, a standard DIC setup measures the test article with a high resolution simultaneously, requiring only a small amount of excitation points, which is conducive to short testing times. Multiple drive points are still required to sufficiently excite all the modes within the bandwidth of interest.

Another metric that is valuable to compare between DIC and accelerometers is temporal resolution or sampling rate which determines the measurable bandwidth. The cameras used for DIC have a maximum frame rate of 20,000 FPS. For this project, the maximum frame rate used was 10,000 FPS yielding a Nyquist frequency of 5000 Hz. The reason that we

used a lower than maximum frame rate was because of lighting limitations. As camera framerate increases, shutter speed decreases, meaning less light is sensed by the camera. This means that at higher framerates, more illumination on the test article is required for the images to be clear enough for displacements to be measured. When we increased the framerate to above 10,000 Hz, our lights were insufficient for our cameras to capture the plate's dynamics.

A significant strength of DIC is the amount of data that it produces. This advantage can also be an obstacle. Especially while all the measurements are in the form of images taken by the cameras, the quantity of data is on the order of hundreds of gigabytes. As a result, the spatial, temporal and frequency resolutions that are possible during a test are all inversely proportional and are limited by the cameras' storage space, which in our case is 64 GB per camera. For example, a test recorded at 10,000 FPS on our cameras allows for 6.55 seconds of recording time. This means that assuming a test will use the cameras' entire memory, and that the spatial resolution is held constant, increasing the bandwidth of a modal test will decrease the frequency resolution.

One limitation of using accelerometers in an impact test is that excessive force can overload the accelerometers. With DIC this limit is removed. The main limit to applied force while using DIC is the force required to overload the force transducer. Additionally, the force should be low enough to not activate nonlinearities or to move the test article out of focus or frame. In our tests, we increased the force amplitude enough to ensure the plate's response would be above the noise floor.

An obstacle that we experienced in our DIC tests is that our setup had no anti-aliasing measures. When analyzing measured frequency responses, we had to question if we were seeing resonances outside of our Nyquist range. One strategy that we used to mitigate this issue was to select impact hammer tips that would excite frequencies within our bandwidth of interest and roll off after that. This method alone was not sufficient because there are only a few options for tip material, which limited control over hammer roll-off.

We were able to identify aliased modes in our data by comparing measured mode shapes that we expected were above Nyquist range to high frequency mode shapes in our FE model. Because of the high spatial resolution of the mode shapes measured by DIC, it was possible to identify features of modes well above our bandwidth of interest. In the high framerate test, we found an aliased mode and were able to identify its true frequency as shown in Figure 9. By knowing this mode was aliased, we could reflect the measured frequency of 4058 Hz about our Nyquist frequency of 5000 Hz to determine that this mode's resonant frequency was approximately 5942 Hz.

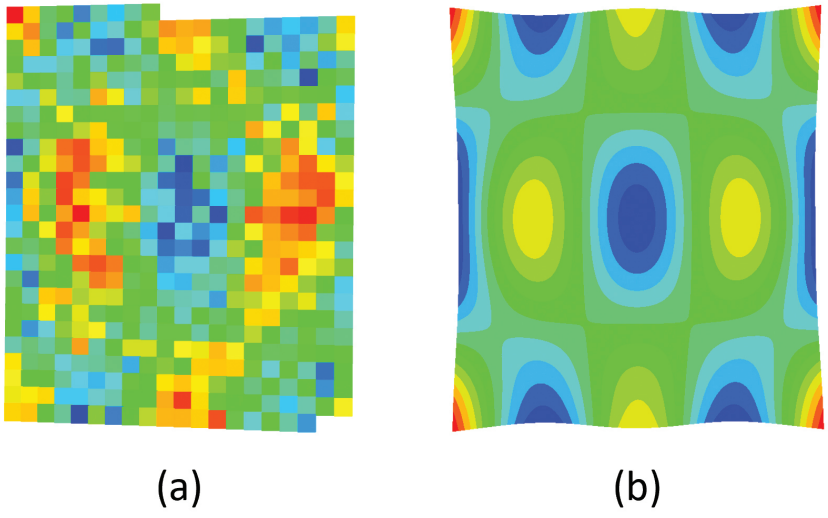


Fig. 9 (a) Aliased mode 18 measured by the 10k fps DIC test at 4058 Hz and (b) mode 18 modeled by the FEA at 5539 Hz. Reflecting about the Nyquist frequency, we measured this mode's resonant frequency to be 5942 Hz

Conclusions

By comparing measurements of impact tests using DIC and accelerometers we made multiple conclusions. We showed that natural frequency measurements of the article with accelerometers attached yield the same results with both measurement methods. Additionally, we qualitatively showed that mode shapes measured with DIC have higher resolution than with

accelerometers but may have more noise. While there was a discrepancy between measurements of damping ratios between the two methods, we showed that DIC measurements have remarkable precision, regardless. This demonstrates that roving hammer impact tests measured with DIC have the potential to perform as well as accelerometers for determining modal parameters, especially at natural frequencies below 5000 Hz.

Furthermore, we showed that attaching accelerometers to our test article led to a significant alteration to the natural frequencies and damping ratios of our test article. We also suspect that attaching accelerometers increased the variability of damping ratio measurements. While DIC measurements without attached accelerometers showed remarkable consistency for all modal parameters, even with varying frame rates, damping ratio measurements didn't appear as consistent when accelerometers were attached. To confirm this conclusion, we would need to perform more DIC and accelerometer measurements of the test article with accelerometers attached. We suspect that external damping related to the wax, cables, and mass loading contributed to a general increase in damping ratios values and their variability.

Future Work

While this paper comparatively analyzes many aspects of DIC significant to its use in modal analysis, there is much opportunity for further research on DIC's advantages and limitations. We recommend testing DIC for modal analysis on test articles with more complex geometries than a rectangular plate. DIC can also be used for applications such as Resonant Ultrasound Spectroscopy, where minimal invasiveness and high-fidelity mode shape characterization is critical. Additionally, we recommend using DIC to quantify strain and modal parameters simultaneously during a high amplitude test for qualification testing.

Acknowledgments We would like to acknowledge the Los Alamos Dynamics Summer School leaders Shannon Danforth and Samantha Ceballes for their leadership and guidance through this project. This research was funded by and performed at Los Alamos National Laboratory (LANL) through the Engineering Institute's Los Alamos Dynamics Summer School. The Engineering Institute is a research and education collaboration between LANL and the University of California San Diego's Jacobs School of Engineering. This collaboration seeks to promote multidisciplinary engineering research that develops and integrates advanced predictive modeling, novel sensing systems, and new developments in information technology to address LANL mission-relevant problems. Los Alamos National Laboratory is an affirmative action/equal opportunity employer, and is operated by Triad National Security, LLC for the National Nuclear Security Administration of U.S. Department of Energy under contract no. 89233218CNA000001. The U.S. Government and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript or allow others to do so, for U.S. Government purposes.

References

1. H. Schreier, J.-J. Orteu, and M. A. Sutton, "Image correlation for shape, motion and deformation measurements: Basic concepts, theory and applications," *Image Correlation for Shape, Motion and Deformation Measurements*, 2009.
2. R. Janeliukstis and X. Chen, "Review of digital image correlation application to large-scale composite structure testing," *Composite Structures*, vol. 271, p. 114143, 2021/09/01/ 2021, doi: https://doi.org/10.1016/j.compstruct.2021.114143.
3. B. Liu et al., "Digital image correlation in extreme conditions," *Thin-Walled Structures*, vol. 205, p. 112589, 2024/12/01/ 2024, doi: https://doi.org/10.1016/j.tws.2024.112589.
4. M. N. Helfrick, C. Niezrecki, P. Avitabile, and T. Schmidt, "3D digital image correlation methods for full-field vibration measurement," (in English), *Mech Syst Signal Pr*, vol. 25, no. 3, pp. 917-927, Apr 2011, doi: 10.1016/j.ymssp.2010.08.013.
5. P. L. Reu, D. P. Rohe, and L. D. Jacobs, "Comparison of DIC and LDV for practical vibration and modal measurements," *Mech Syst Signal Pr*, vol. 86, pp. 2-16, 2017/03/01/ 2017, doi: https://doi.org/10.1016/j.ymssp.2016.02.006.
6. C. Warren, C. Niezrecki, P. Avitabile, and P. Pingle, "Comparison of FRF measurements and mode shapes determined using optically image based, laser, and accelerometer measurements," *Mech Syst Signal Pr*, vol. 25, no. 6, pp. 2191-2202, 2011/08/01/ 2011, doi: https://doi.org/10.1016/j.ymssp.2011.01.018.
7. J. Javh, J. Slavič, and M. Boltežar, "High frequency modal identification on noisy high-speed camera data," *Mech Syst Signal Pr*, vol. 98, pp. 344-351, 2018/01/01/ 2018, doi: https://doi.org/10.1016/j.ymssp.2017.05.008.
8. J. E. T. Penny, M. I. Friswell, and S. D. Garvey, "Automatic choice of measurement locations for dynamic testing," *AIAA Journal*, vol. 32, no. 2, pp. 407-414, 1994, doi: 10.2514/3.11998.
9. C. W. Miller, "Determination of Modal Parameters from Experimental Frequency Response Data," 1978.

Chapter 5
Capturing 3D Mode Shapes with Cell Phone Videos

Shawn Richardson and Mark Richardson

Abstract Everyone has a cell phone in their pocket which has a video camera in it. A cell phone video is a low-cost, non-contacting way of recording the vibration of rotating machinery and the resonant vibration of mechanical structures. A cell phone can be used as a quick way to visualize vibration and complement other **ODS** and **Modal** testing methods

For this paper, **ODS-FRF** measurement functions were calculated from time waveforms (**TWFs**) that were extracted from cell phone videos of a rotating machine. Then the **ODS-FRFs** were curve fit using **FRF-based** curve fitting to yield **2D OMA** mode shapes of the machine. Two sets of **2D OMA** mode shapes were extracted from videos recorded at ***right angle views*** of the machine, one from the **Side** and one from the **Top**.

Then, **3D OMA** mode shapes were created by adding the mode shape components from the missing third direction to the **2D OMA** mode shapes from each view. The only requirement for preserving the phases of the missing mode shape components is that the **ODS-FRFs** share a common reference response.

Keywords Time Waveform (**TWF**) · Digital Fourier Transform (**DFT**), Auto Power Spectrum (**APS**) · Cross Power Spectrum (**XPS**).**ODS-FRF** · Operational Modal Analysis (**OMA**) mode shape

Rotating Machine

The test article used for this paper is shown on Figure 1. The machine speed was fixed at ***approximately* 1000 RPM** throughout all the cell phone video recordings.

For this paper, all cell phone videos were recorded with one screw added to both the inner and outer rotor. Previous papers [1, 2] have presented results when different unbalancing screws were used to a similar rotating machine.

Introduction

Using accelerometers for data acquisition is more expensive and time-consuming to implement when compared to recording a cell phone video. Furthermore, because it is non-contacting, a cell phone can record vibration of machinery or structural parts that are too hot, too dangerous, or are inaccessible for attaching accelerometers.

TWFs & DFTs

In previous papers [1] to [4], traditional signal processing methods were applied to the **TWFs** extracted from the frames of a video. **TWFs** can be used to display **time-based ODSs** on a point grid, allowing vibration to be visualized at slower speeds with higher amplitudes. And their associated **DFTs** can be used to display **frequency-based ODS**'s using sinusoidal animation of the **ODS** at the cursor location in the **DFTs**.

When a video is processed in MEscope [10], a rectangular grid of points with surfaces between them is created. Video frames are attached to this surface during animated display of Video ODS's. Using a point grid, millions of pixels per frame

Shawn Richardson and Mark Richardson
Vibrant Technology, Inc., Centennial, CO
e-mail: shawn.richardson@vibetech.com; mark.richardson@vibetech.com

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

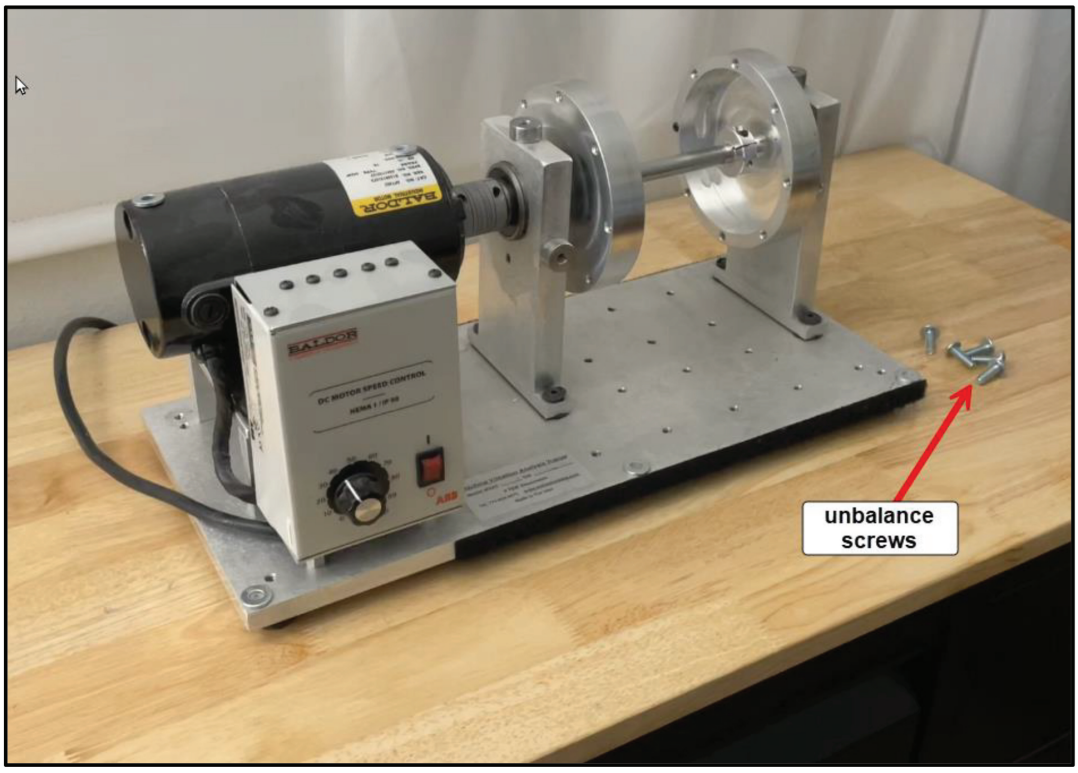


Fig. 1 Rotating Machine Showing Unbalance Screws Added to Its Rotors

of a cell phone video are processed to extract TWFs for the horizontal & vertical deflections of thousands of points in the point grid. Grid points with little or no motion, (e.g. background points), are hidden and their linked TWFs removed from further analysis.

A DFT is calculated from each TWF that is extracted from the video. Time-based ODS's are displayed in animation from the TWFs using a sweeping Line cursor. Frequency-based ODS's are displayed in animation from the DFTs using sinusoidal dwell modulation of the ODS at a cursor position. The magnitude & phase of the ODS at selected grid points can also be displayed, as shown in Figures 2 & 3.

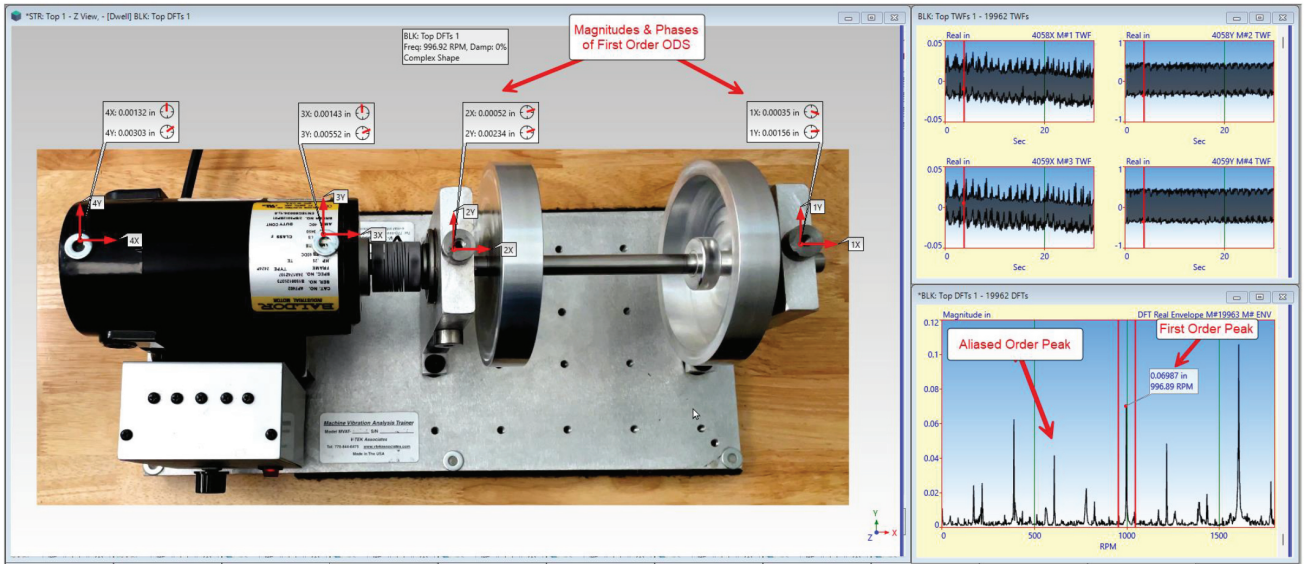


Fig. 2 Top View-First-Order ODS Animated from DFTs

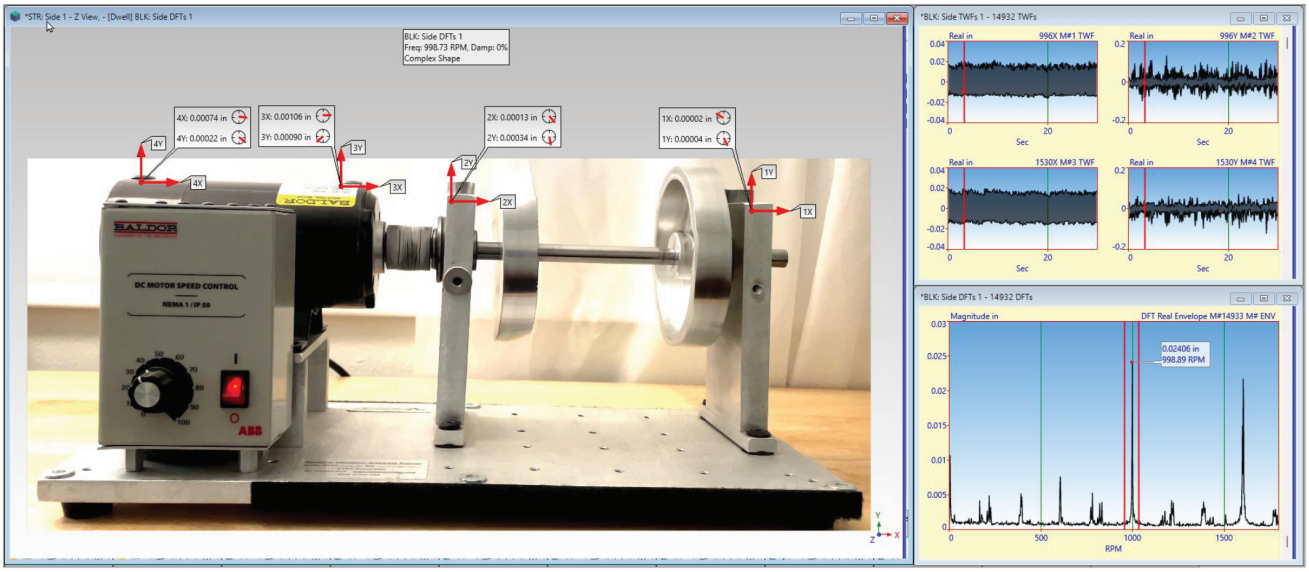


Fig. 3 Side View-First-Order ODS Animated from DFTs

ODS-FRFs

A unique frequency domain function, called an **ODS-FRF,** can be calculated from each response **TWF** extracted from a video. A set of **ODS-FRFs** calculated from the **TWFs** is typically more accurate than **DFTs** because ***spectrum averaging*** can be used to reduce extraneous noise from the **ODS-FRFs**, and **Hanning** windowing can be used to ***reduce signal leakage*** from the resonance peaks.

The ***magnitude*** of an **ODS-FRF** is the **APS** of a ***response*** **DOF** at a grid point. The ***phase*** of the **ODS-FRF** is the ***phase of the*** **XPS** between the ***response*** **DOF** and the **DOF** of any ***reference*** grid point.

ODS-FRFs carry the ***same units*** as the response **TWFs** from which they are calculated. Unlike an **XPS**, an **ODS-FRF** in displacement units can be ***accurately differentiated to velocity units*** by multiplying it by the frequency variable.

Vibration in ***velocity units*** is commonly used by vibration analysts to quantify vibration levels in rotating equipment.

One of the laws of the **FFT** algorithm is that $\mathbf{\Delta f = 1/T}$, where $\mathbf{\Delta f}$ is the ***frequency difference between samples*** of an **ODS-FRF**, and **T** is the ***time length*** of **TWF** data from which the **ODS-FRF** was calculated. For example, if an **ODS-FRF** is calculated from a **TWF** using **25 second**s of uniformly sampled data, the frequency resolution ($\mathbf{\Delta f}$) of the **ODS-FRF** is **60/25 = 2.4 RPM**.

To increase the frequency resolution of an **ODS-FRF**, **TWF** data over a ***longer period*** **T** is required. Therefore, the video from which the **TWFs** are extracted must be recorded over a ***longer period*** **T**.

Aliased Order Peaks

A limitation of any video recording is that ***anti-alias filtering cannot be used*** to remove high-frequency signals from the video.

Without anti-alias filtering, machine order peaks ***at frequencies greater than one-half the sampling frequency***, (called **Fmax)**, are ***folded around*** **Fmax** and appear at lower frequencies in the **ODS-FRFs**, as shown in Figures 2, 3, & 4.

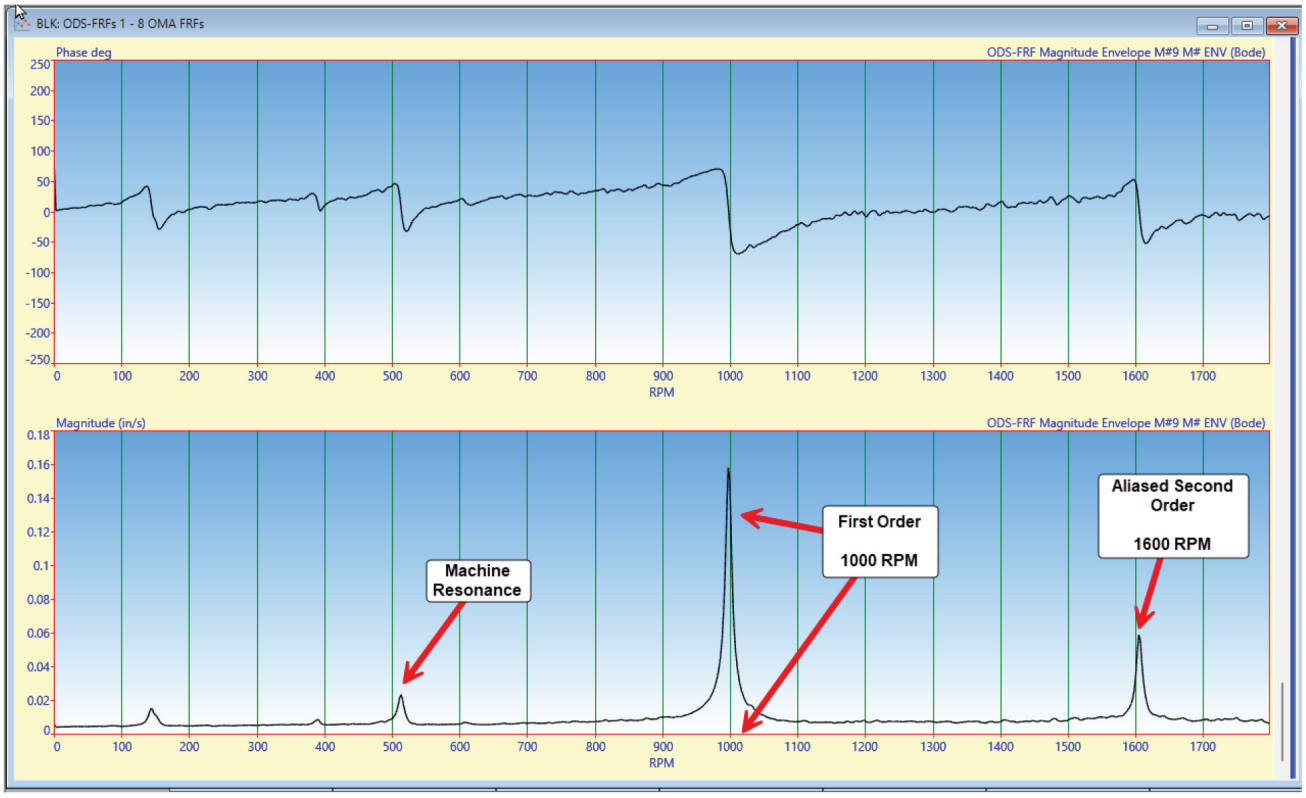


Fig. 4 ODS-FRF Showing First Order, Aliased Second & Machine Resonance

All machine-order resonance peaks between **Fmax & 2x Fmax** are folded around (***wrapped around***) **Fmax** and appear at lower frequencies in the frequency band **(0 to Fmax)**. Aliasing of higher frequencies occurs in both **DFTs** and **ODS-FRFs**.

The magnitude of the **ODS-FRF** shown in Figure 4 was calculated from a **TWF** that was extracted from a video that was sampled at **60 fps**, or **3600 RPM**. Therefore, the **Fmax** of the **ODS-FRF** is **1800 RPM**. The first-order peak is at the machine running speed and is clearly visible at **approximately 1000 RPM**. The second order resonance peak should be at about **2000 RPM** and third order resonance peak should be at about **3000 RPM, etc.** Aliased order peaks are clearly visible ***at lower frequencies*** in the **DFTs** and the **ODS-FRFs**.

The **ODS** or **OMA** mode shape of an aliased order is correct. Only its frequency is incorrect.

The aliased frequency of an order higher than **Fmax** can be calculated from **Fmax** and the expected order frequency.

- **Second order** aliased frequency → 1800 - (2000 - 1800) **1600 RPM**
- **Third order** aliased frequency → 1800 - (3000 - 1800) **600 RPM**
- **Fourth order** aliased frequency → 1800 –((3600 - (4000 - 3600)) -1800) **400 RPM**

The **ODS-FRF** in Figure 4 was calculated from a **TWF** with ***a 25-second length***. Therefore, **Δf = (1/25) Hz or 60/25 = 2.4 RPM** and the aliased frequency of the third order peak is ***within one*** **(Δf)** of its calculated value.

Curve Fitting the ODS-FRFS

ODS-FRFs can be curve-fit using an **FRF-based** curve fitter if they have been filtered with a special window. In MEscope [10] this filter is called a ***DeConvolution window***. This filter reshapes an **ODS-FRF** so that it resembles an **FRF** waveform and therefore can be curve fit using **FRF-based** curve fitting. An **FRF-based** curve fit of an **ODS-FRF** is shown in Figure 5. In the upper left-hand graph, the **red curve fitting function** is overlaid on a magnitude plot of the **ODS-FRF**.

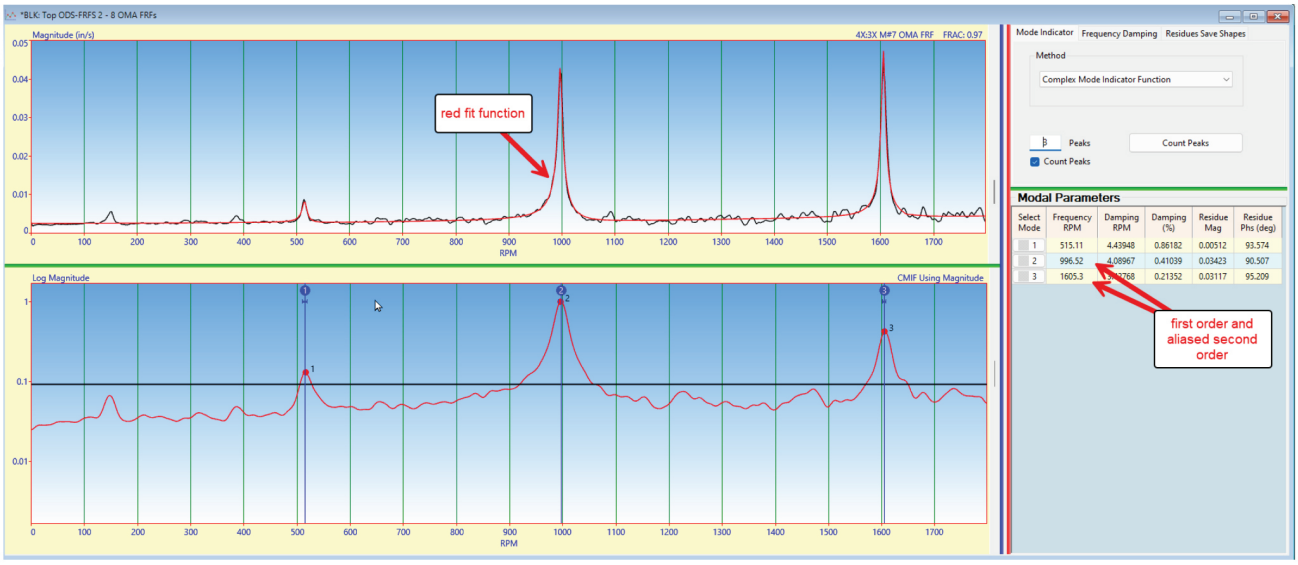


Fig. 5 Curve Fit of ODS-FRFs.

Comparing Mode Shapes from Successive Side View Videos

Two successive video recordings from a **Top** view of the rotating machine were post-processed and the resulting mode shapes were compared using two different shape correlation coefficients, the Modal Assurance Criterion (**MAC**) and the Shape Difference Indicator (**SDI**). Both correlation coefficients have values between 0 and 1. A value of **"0"** means that two shapes are different from one another, and a value of **"1"** means that two mode shapes are the same.

Each shape has 8 components corresponding to X & Y deflections at 4 grid points. Two points are on the motor, and one is on the top of each bearing block. The **MAC** & **SDI** values between the **OMA** mode shapes from two successive videos of the machine are shown on Figures 6 & 7.

In Figure 6, mode shapes 2 & 3 correlate well from two successive **Side** view videos, with **MAC** & **SDI** values above 90% In Figure 7, all three mode shapes correlate well from two successive **Top** view videos, with **MAC** & **SDI** values above 90%.

The low values of the off-diagonal MAC & SDI values indicate that the three **OMA** mode shapes have different magnitudes & phases. This becomes clear when the magnitudes & phases of the mode shapes are displayed on the point grid.

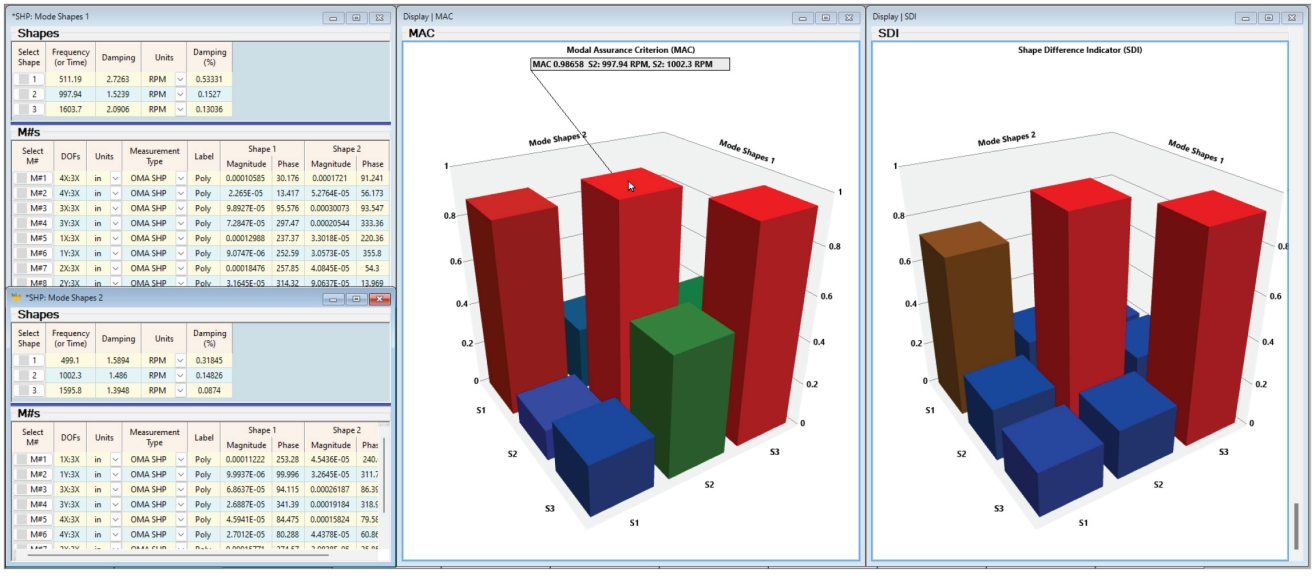


Fig. 6 MAC & SDI of Mode Shapes From Two Sucessive Side Videos

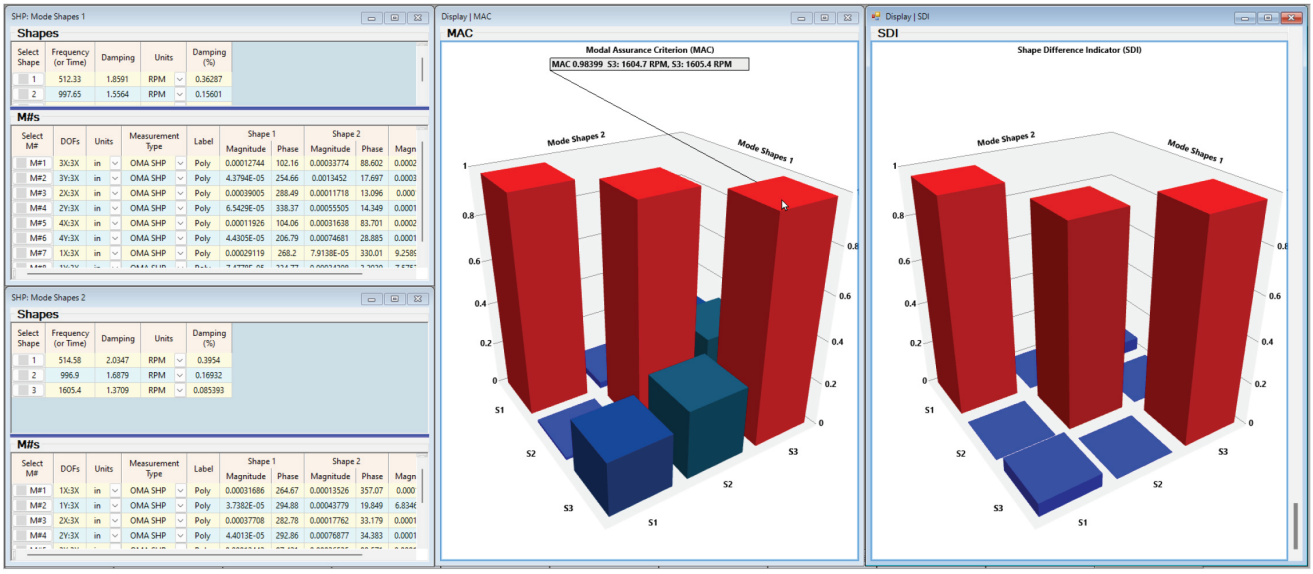


Fig. 7 MAC & SDI of Mode Shapes From Two Sucessive Top Videos

Comparing Mode Shapes from Successive Top View Videos

Two successive video recordings from a side view of the rotating machine were post-processed and the resulting mode shapes are compared the Modal Assurance Criterion (MAC) and the Shape Difference Indicator (SDI) in Figure 6.

Side View Animated Display of the 3D OMA Mode Shapes

Animation of the **3D OMA** mode shape of the first order at **997 RPM** from the **Side** view is shown in Figure 8. In the X-direction the motor and inner bearing block are in-phase, and the outer bearing block is 180 degrees out-of-phase. In the Y-direction, the front side of the motor is in-phase with both bearing blocks and the back side is 90 degrees out-of-phase. In the Z-direction, the motor and bearing blocks are all in-phase with one another.

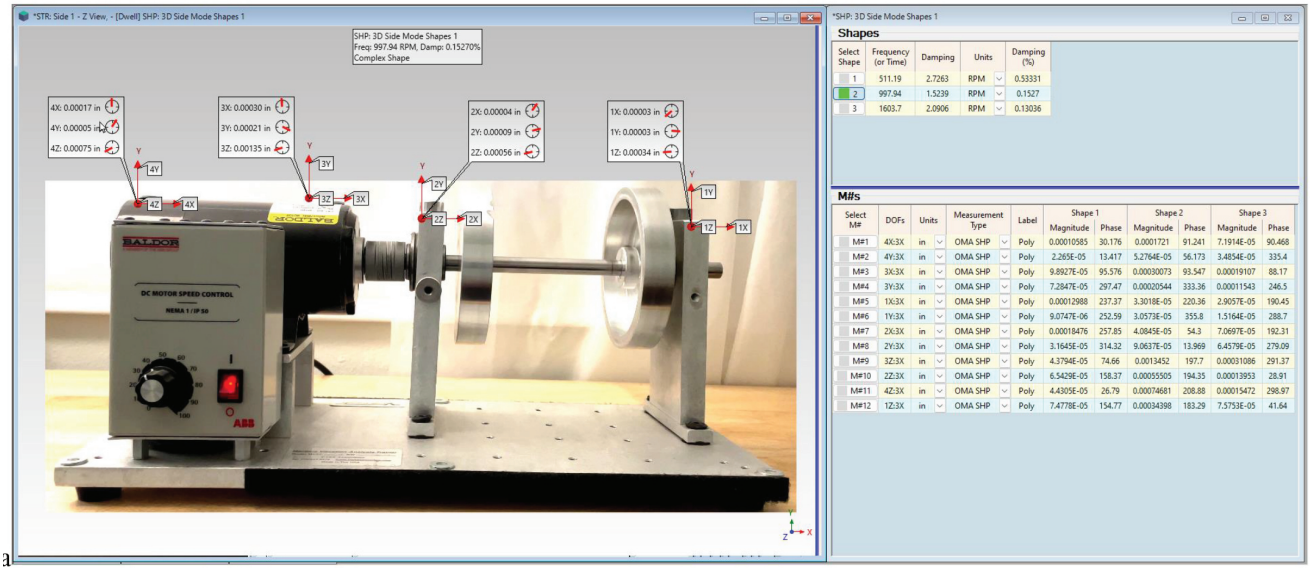


Fig. 8 Side View - Animated Deflection of the First-Order OMA Mode Shape

Top View Animated Display of the 3D OMA Mode Shapes

Animation of the **3D OMA** mode shape of the first order at **997 RPM** is shown in Figure 9. In the X-direction the motor is 90 degrees out-of-phase with the bearing blocks. In the Y-direction, the motor and bearing blocks are in-phase. This agrees with the Z-direction in the Side view. In the Z-direction, the front side of the motor is in-phase with both bearing blocks and the back side is out-of-phase. This agrees with the Y-direction in the Side view.

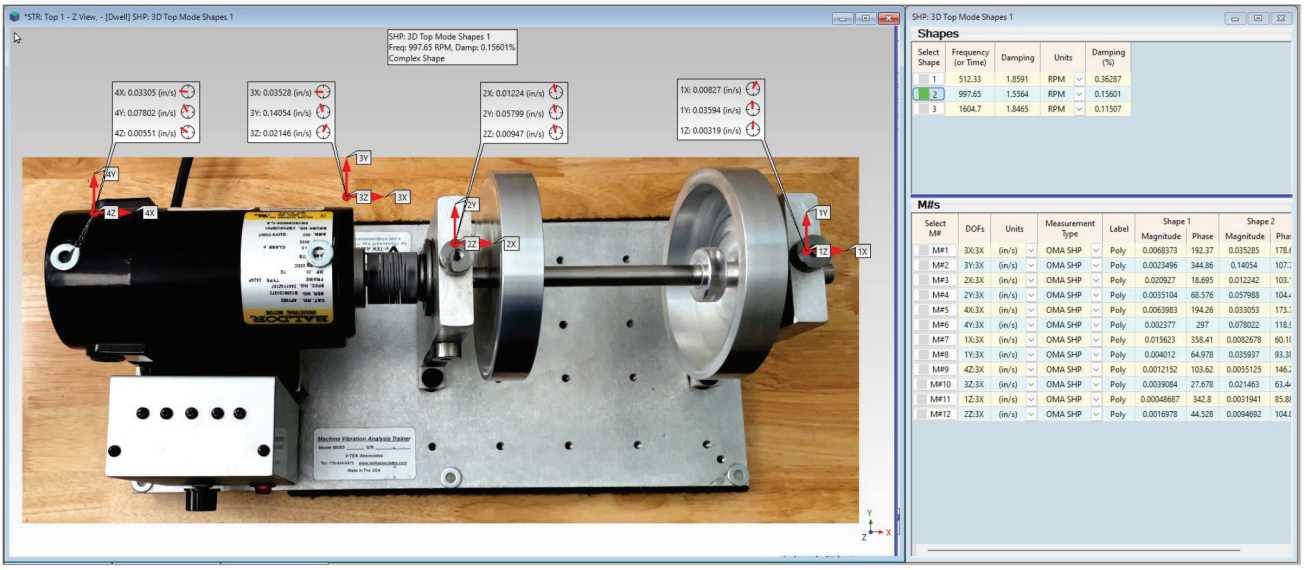


Fig. 9 Top View - Animated Deflection of the First-Order OMA Mode Shape

Conclusions

In this paper, we extracted **2D** mode shapes from two successive videos of the **Top** view and **Side** view of a rotating machine while it was running at approximately **1000 RPM**. Then, **ODS-FRFs** were calculated from the **TWFs** using **25-seconds** of data per average, 5 spectrum averages with 95% overlap processing and a Hanning window to reduce signal leakage from the resonance peaks. Then the **ODS-FRFs** were windowed with a Deconvolution window to prepare them for **FRF-based** curve fitting. **2D OMA** mode shapes of a rotating machine were obtained by applying **FRF-based** curve fitting to the **ODS-FRFs**.

The **ODS-FRFs** contained three peaks, the running speed at **1000 RPM**, the aliased second order at **1800 RPM**, and a machine resonance at **500 RPM**.

The **2D OMA** mode shapes obtained from curve fitting **ODS-FRFs** calculated from successive videos of the Top and Side of the machine were correlated using the Modal Assurance Criterion (**MAC**) and Shape Difference Indicator (**SDI**). Mode shape pairs with coefficients above 90% correlated well.

Finally, **3D OMA** mode shapes of the machine were created by adding the Y-direction mode shape components from the **Top** view as Z-direction components to the mode shapes of the **Side** view. Likewise, the Y-direction mode shape components from the **Side** view were added as Z-direction components to the mode shapes of the **Top** view.

All mode shape phases were preserved by using **DOF 3X** as the reference for the **ODS-FRF** calculations. No tachometer signal of other means of matching phases is needed when **ODS-FRFs** are calculated using the same reference response **DOF**.

References

1. S. Richardson, M. Richardson “Using Mode Shapes From Cell Phone Videos For Machinery Health Monitoring” VIATC 2025 Newport News, VA August 6-8, 2025
2. S. Richardson, M. Richardson “Extracting 3D Mode Shapes from a Cell Phone Video” IMAC XLII Orlando, FL February 10-13, 2025

3. D. Ambre, B. Schwarz, S. Richardson, M. Richardson, “Using Cell Phone Videos to Diagnose Machinery Faults” IMAC XLI Austin, TX February 13-16 2023
4. B. Schwarz, S. Richardson, M. Richardson, “Using a Cell Phone Video and ODS Correlation to Diagnose Unbalance in Rotating Machinery” IMAC XLII January 29 - February 1, 2024 Orlando, FL
5. M.H. Richardson, “Is It a Mode Shape or an Operating Deflection Shape?” Sound and Vibration magazine March 1997
6. B. Schwarz, M.H. Richardson, “Measurements Required for Displaying Operating Deflection Shapes” Proceedings of IMAC XXII January 26, 2004
7. B. Schwarz, M.H. Richardson, “Introduction to Operating Deflection Shapes” CSI Reliability Week Orlando, FL October 1999
8. S. Richardson, J. Tyler, P. McHargue, M. Richardson “A New Measure of Shape Difference” IMAC XXXII February 3-6, 2014
9. S. Richardson, J. Tyler, B. Schwarz, P. McHargue, M. Richardson “Using Modal Parameters for Structural Health Monitoring” IMAC-XXXV Garden Grove, CA January 30–February 2, 2017
10. MEscope^{TM} is a trademark of Vibrant Technology, Inc. www.vibetech.com

Chapter 6
Streaming PCA for Real-Time Structural Dynamics Tracking in Event-Based Vision

Moises Felipe Silva, Anthony Starfleaf, Andre Green, Nicholas Bruns, and David Mascarenas

Abstract Event-based neuromorphic imagers deliver asynchronous streams of pixel events encoding scene dynamics with microsecond resolution, but extracting precise object trajectories from such high-throughput data remains challenging. We introduce a set of PCA-based tracking methods that use the dominant direction of recent events to find and follow moving targets. Each method keeps a short history of event coordinates and updates its estimated principal direction of motion - either by recalculating from scratch or by fast incremental updates - so that the many event points collapse into a simple, low-dimensional view. In this view, each object appears as a clear feature, which we locate, refine to get precise positions, and then link over time to form continuous tracks. We demonstrate these approaches on scenes with several objects moving at different speeds and rotational rates, showing they offer reliable, scalable multi-object tracking with minimal computation. These PCA-driven techniques form a lightweight, flexible toolkit for real-time motion analysis in robotics, surveillance, structural monitoring, and other applications using event-based vision.

Keywords event-based cameras · machine learning · motion tracking · PCA · low-light sensing, tracking · object tracking

Introduction

Traditional cameras capture scenes as a sequence of complete frames, which imposes both latency and redundancy. In contrast, neuromorphic or event-based sensors operate by detecting and reporting changes in brightness at each pixel as they occur. This asynchronous design delivers temporal resolutions on the order of microseconds, effective rates well above 5,000 frames per second, and dynamic ranges exceeding 120 dB - greater than those available from conventional imagers [1, 2]. The output of such sensors resembles the operation of biological retinas, hence their frequent description as silicon retinas.

Despite these advantages, much of the current practice in event-based vision involves reassembling streams of events into artificial frame-like structures. While convenient, this approach diminishes the unique benefits of event data, especially its ability to encode motion directly at very fine time scales [3]. A more natural interpretation is to regard the event stream as a high-dimensional point cloud evolving in space and time. Within this representation, statistical tools such as principal component analysis (PCA) are well suited to expose dominant directions of motion and to reduce data complexity [1].

Here, we advance this perspective by formulating a tracking method that operates entirely in event space. The proposed algorithm maintains a continuously updated PCA representation, either on an event-by-event basis for minimal latency, or over short event buffers to improve throughput. These representations serve as the foundation for efficient object localization and, by extending the centroid trajectories into the frequency domain, for real-time spectral estimation as well. We evaluate the framework on synthetic and laboratory data of oscillating objects, demonstrating accurate recovery of both motion and oscillation characteristics at low computational cost.

Moises Felipe Silva(✉) · Anthony Starfleaf · Andre Green · Nicholas Bruns · David Mascarenas
Los Alamos National Laboratory, Los Alamos, NM 87545
e-mail: mfelipe@lanl.gov; aglor@lanl.gov; andre_green@lanl.gov; bruns@lanl.gov; dmascarenas@lanl.gov

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

Related Works

Efforts to harness the capabilities of event-based sensors span both adaptations of existing vision pipelines and new algorithms designed for asynchronous data. One of the earliest strategies was to accumulate events into frame-like images or voxel grids, thereby enabling the reuse of conventional vision algorithms. While effective for compatibility, this practice discards much of the fine temporal structure that makes event data attractive in the first place [3].

Recent approaches emphasize processing events directly, with PCA standing out for such applications. Khairallah et al. [4] demonstrated that local principal components could be used to infer optical flow vectors for odometry at real-time rates, while Ramesh et al. [5] introduced PCA-RECT, which applied PCA to normalized event patches for efficient object detection on FPGA platforms. These works illustrate PCA's efficiency but are largely confined to single-object motion without addressing concurrent frequency recovery. Broader tracking methods have been explored by Hamann et al. [6], where a timestamp-driven scheme for following arbitrary features across the sensor has shown robustness under extreme speed and lighting conditions. On the other hand, deep learning approaches have exploited event sparsity using transformer and attention mechanisms [7], achieving impressive accuracy but often at high computational cost.

On a different note, another important marker of a subject's characterization is its oscillatory frequency content to be estimated from event streams. Pfrommer et al. [8] showed that flicker frequencies can be measured directly from asynchronous pixel activity, and later simulation studies investigated how different motion regimes affect frequency estimation [9]. Multi-window ensemble approaches have also been applied to location tracking, balancing the trade-off between latency and robustness [10]. These methods demonstrate the potential of using event data for spectral analysis, though they typically consider single or static sources rather than multiple moving targets.

Therefore, our work builds upon these foundations by integrating localization, trajectory recovery, and spectral estimation into a single PCA-based framework. Unlike earlier PCA applications, our methods maintain streaming updates, and operate effectively in multi-object scenarios. This positions the approach as a bridge between statistical PCA techniques and the broader goals of event-driven spectral and motion analysis.

Background

Problem Statement

Considering each asynchronous event as

$$\boldsymbol{e_i} = (\boldsymbol{x_i}, \boldsymbol{y_i}, \boldsymbol{t_i}),$$

where (x_i, y_i) is the pixel location and t_i the i-th timestamp. We group events into non-overlapping windows of duration T (per-buffer mode) or into a sliding window of duration T that advances by each new event (per-sample mode). Then, for the k-th window, we collect the N_k events into a data matrix

$$X^k = \begin{bmatrix} x_1 & x_2 & \cdots & x_{N_k} \\ y_1 & y_2 & \cdots & y_{N_k} \end{bmatrix} \in \mathbb{R}^{2\times N_k}.$$

We seek a rank-r subspace, spanned by orthonormal basis $W \in \mathbb{R}^{2\times r}$, that best reconstructs the centered data set X. That is analogous to the PCA minimization function via the eigen decomposition of the covariance matrix [11]:

$$C^k = \frac{1}{N_k}\left(X^k - \mu 1^\intercal\right)\left(X^k - \mu 1^\intercal\right)^\intercal,$$

where $\mu \in \mathbb{R}^2$ is the sample mean and 1 is a length-N_k vector of ones for padding. Solving this yields the top r eigenvectors of C^r, which become the columns of W. In per-sample mode, these statistics are updated with each new event, while in per-buffer mode, they are recomputed over the entire batch of N_k events at once. Then, for the i-th event, we define its running means as:

$$\mu_x = \frac{1}{N_k}\sum_{i=1}^{N_k} x_i, \qquad \mu_y = \frac{1}{N_k}\sum_{i=1}^{N_k} y_i,$$

and the centered covariances as

$$Cov\,(X,X) = \frac{1}{N_k - 1}\sum_{i=1}^{N_k}(x_i - \mu_x)^2, \qquad Cov\,(Y,Y) = \frac{1}{N_k - 1}\sum_{i=1}^{N_k}(y_i - \mu_y)^2,$$

where the full multivariate covariance matrix becomes

$$Cov\left(X,Y\right)=\frac{1}{N_k-1}\sum_{i=1}^{N_k}\left(x_i-\mu_x\right)\left(y_i-\mu_y\right),$$

deriving a symmetric and positive semidefinite matrix. Therefore, solving the eigenproblem $\det\left(C-\lambda I\right)\ =0$ yields the characteristic equation

$$\lambda^2-\left(Cov\left(X,\ X\right)+Cov\left(Y,\ Y\right)\right)\lambda+\left[Cov\left(X,\ X\right)+Cov\left(Y,\ Y\right)-Cov(X,\ Y)^2\right]=0,$$

whose roots

$$\lambda_{1,2}=\frac{Cov\left(X,\ X\right)+Cov\left(Y,\ Y\right)\pm\sqrt{\left(Cov\left(X,\ X\right)-Cov\left(Y,\ Y\right)\right)^2+4\,Cov(X,Y)^2}}{2}$$

are the principal variances and the corresponding unnormalized eigenvectors admit the closed form

$$v_i\propto\left[\begin{array}{c}\lambda_i-Cov\left(Y,Y\right)\\Cov\left(X,Y\right)\end{array}\right],\ \ i=1,2.$$

After the variances have been normalized to unit length, projecting onto $[v_1,v_2]$ yields the best rank-2 reconstruction of the event cloud. To avoid $\mathcal{O}(N_k)$ recomputation on each event, when a new event (x_{new},y_{new}) arrives and the oldest (x_{old},y_{old}) is evicted, we update the means in constant time:

$$\mu_{x_{new}}\ =\ \mu_x\ +\frac{x_{new}\ -\ x_{old}}{N_k},\quad \mu_{y_{new}}=\ \mu_y\ +\frac{y_{new}\ -\ y_{old}}{N_k}.$$

In a similar fashion, each covariance entry is thus updated by removing the outgoing contribution and adding the incoming one:

$$Cov(X,Y)_{new}\ =\ Cov\left(X,Y\right)+\frac{\left[\left(x_{new}\ -\ \mu_{x_{new}}\right)\left(y_{new}\ -\ \mu_{y_{new}}\right)-\ \left(x_{old}\ -\ \mu_x\right)\left(y_{old}\ -\ \mu_y\right)\right]}{\left(N_k-1\right)}.$$

Finally, we recompute $\lambda_{1,2}$ and $v_{1,2}$ via the closed-form formulas above, achieving an exact sliding-window PCA in constant time per event. This whole process is also illustrated by Figure 1.

To estimate the positions of multiple moving objects, we project the events onto the principal axis identified by PCA. Each event is represented by a scalar coordinate along this axis, producing a one-dimensional distribution of activity. This distribution is discretized into equally spaced bins, forming a running histogram of event counts that evolves with the sliding buffer. The projection of each event onto the principal axis is given by

$$s_i=v_1^{\intercal}\left[\begin{array}{c}x_i\\y_i\end{array}\right].$$

For the histogram, we define a set of bin edges with spacing γ, covering the full range of projected coordinates. Each incoming event is assigned to one of these bins according to its projected value:

$$\left[s_{min}\left(j-1\right)\gamma,s_{min}+j\,\gamma\right),\ \ j=1,\ldots,\ \gamma.$$

As the buffer advances, updates are handled incrementally. When an old event leaves the buffer, the count in its corresponding bin is decreased, and when a new event enters, the count of its bin is increased. This allows the histogram to be maintained efficiently without recomputation. From this histogram, coarse positions of objects are obtained by identifying peaks with the highest event counts. To improve accuracy, we refine each peak location by computing an average-smoothed centroid over a neighborhood of bins within a half-width γ. This refinement produces a sub-bin resolution estimate of each object's location.

Once positions have been estimated over time, the resulting trajectories are transformed into the frequency domain using a discrete Fourier transform. The dominant peak in the frequency spectrum corresponds to the oscillatory behavior of the tracked object, and by mapping this peak to the physical sampling interval, we recover its oscillation frequency. The histogram-based object's localization is summarized in Figure 2. These steps altogether allow to maintain a high-resolution spatial histogram of events along the principal PCA axis, enabling efficient position estimation at event-rate.

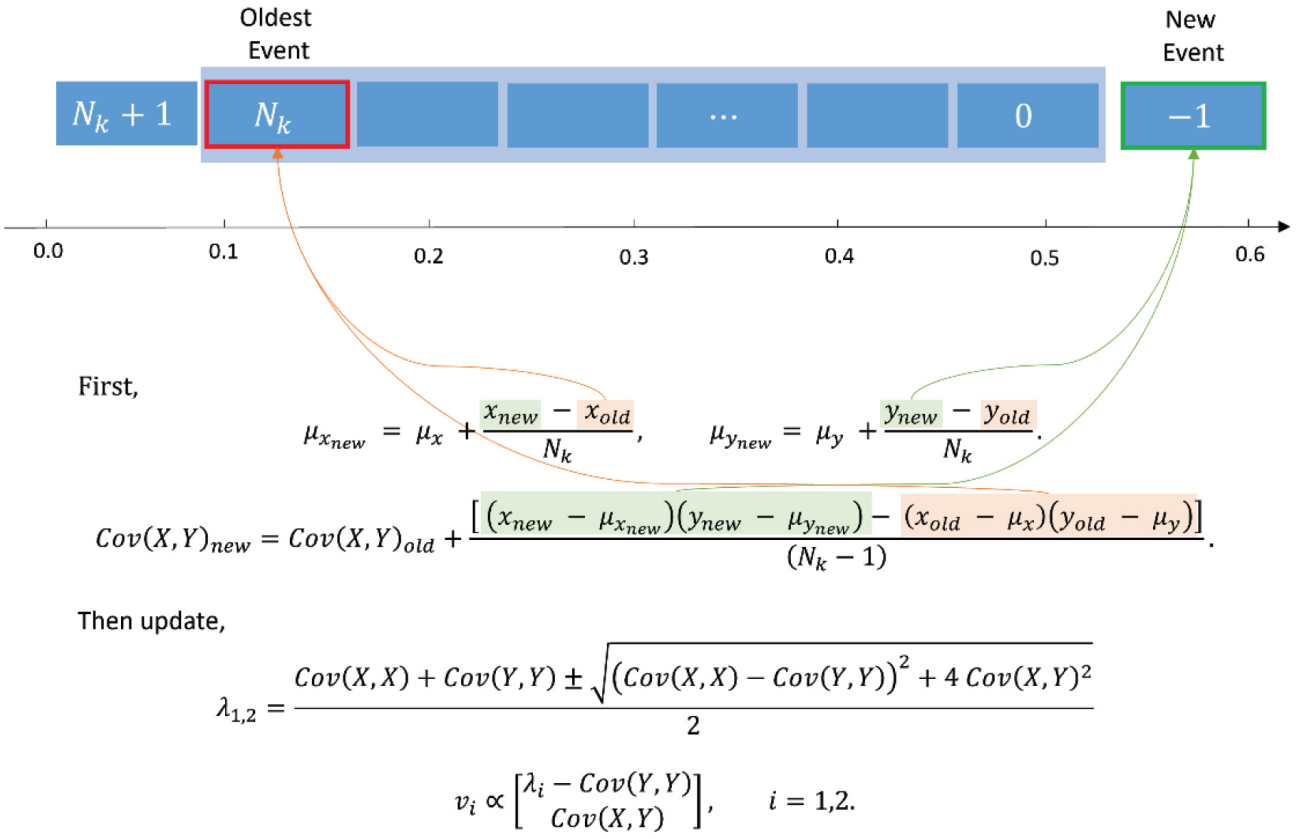


Fig. 1 Sliding-window PCA update. The buffer (blue) shifts forward by dropping the oldest event (red) and adding the newest (green). Running means and covariances are updated incrementally, enabling constant-time PCA without full computation.

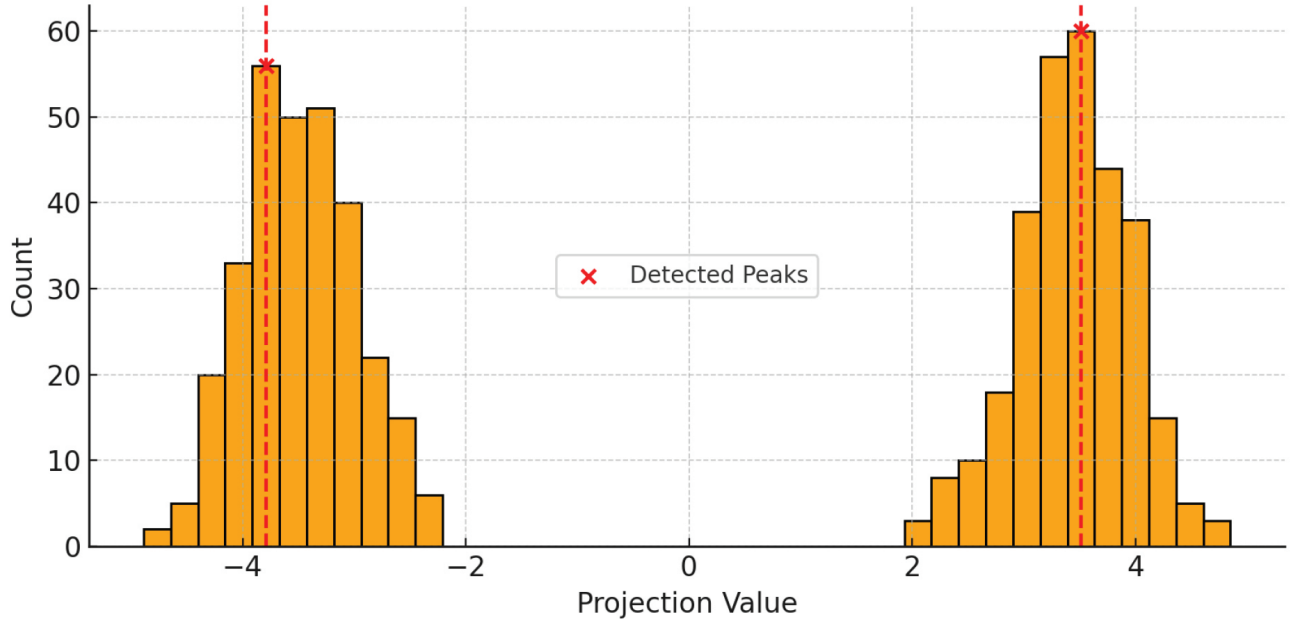


Fig. 2 Object localization along the PCA axis. Events are projected onto the dominant direction, binned into histograms (orange), and the peak bin is refined with a local centroid estimate (red) for precise position recovery.

Results

Simulated Experiments

To validate the proposed framework, we conducted a series of simulated experiments. We modeled moving objects as a sparse set of Gaussian discrete points that both oscillate at a prescribed frequency and may or may not move following a circular trajectory. The sensor field is defined by horizontal and vertical bounds of 240×180 px, simulating conventional event-based imagers. Therefore, the center of motion is taken as the midpoint of those ranges. In the case of N moving targets, their orbit radii are chosen by linearly spacing values between one-eighth and one-third of the field width, and we pick rotation speeds evenly between 0.1 and 0.2 revolutions per second. Each moving object also receives a random phase offset in $[0,\ 2\pi)$ to avoid synchronous trajectories.

At each time step, for each object, we define an oscillation (blinking) cycle to have it outputting "on" or "off" events. In that case, the first half of the cycle the object emits "on" events and output "off" events for the other half of the cycle. When "on", it generates a set number of points spread around its current position in a soft, cloud-like pattern defined by a Gaussian distribution centered at $(c_x,\ c_y)$, where

$$c_x = x + r\cos(\omega t + \phi)\ ,\ c_y = y + r\sin(\omega t + \phi)\ .$$

To emulate background and sensor noise, we inject a Poisson-like cloud of uniformly distributed noise events across the full field of view of the simulated sensor. Figure 3 shows four different simulated scenarios: in (a), three stationary objects blink at distinct frequencies; in (b–d), the same objects move along different trajectories while blinking at varying rates.

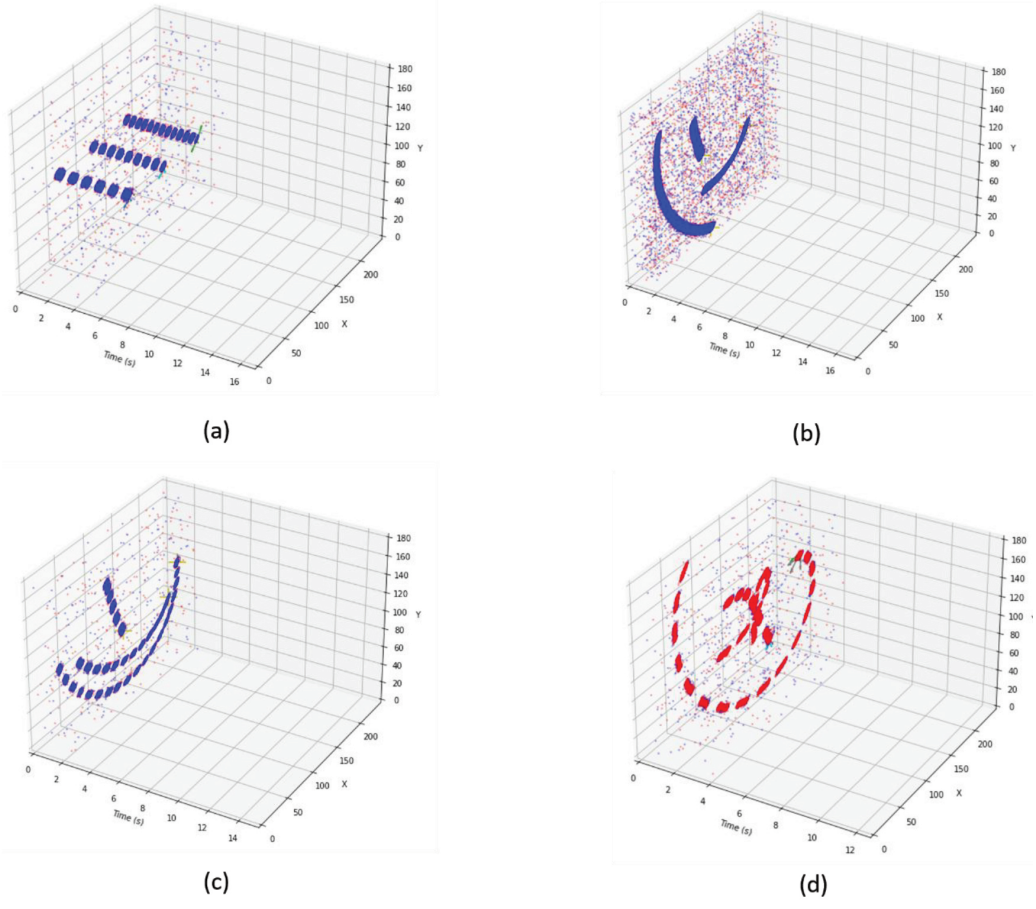
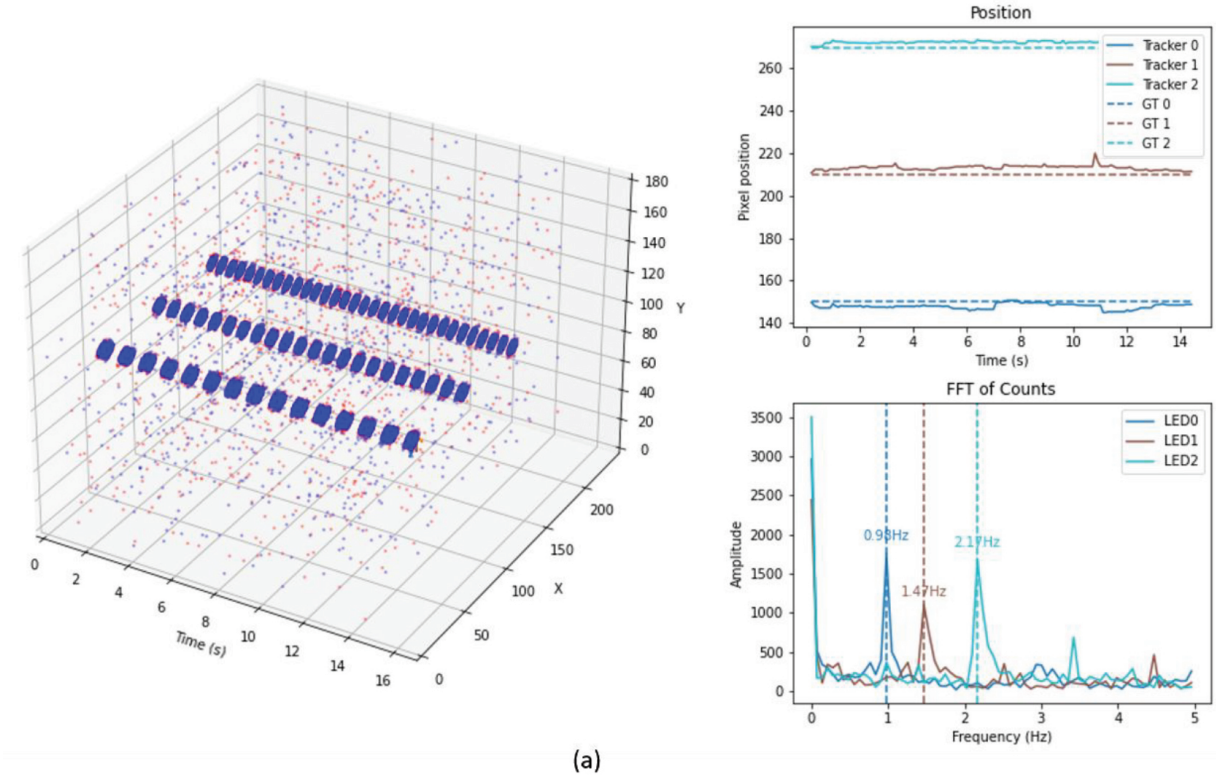
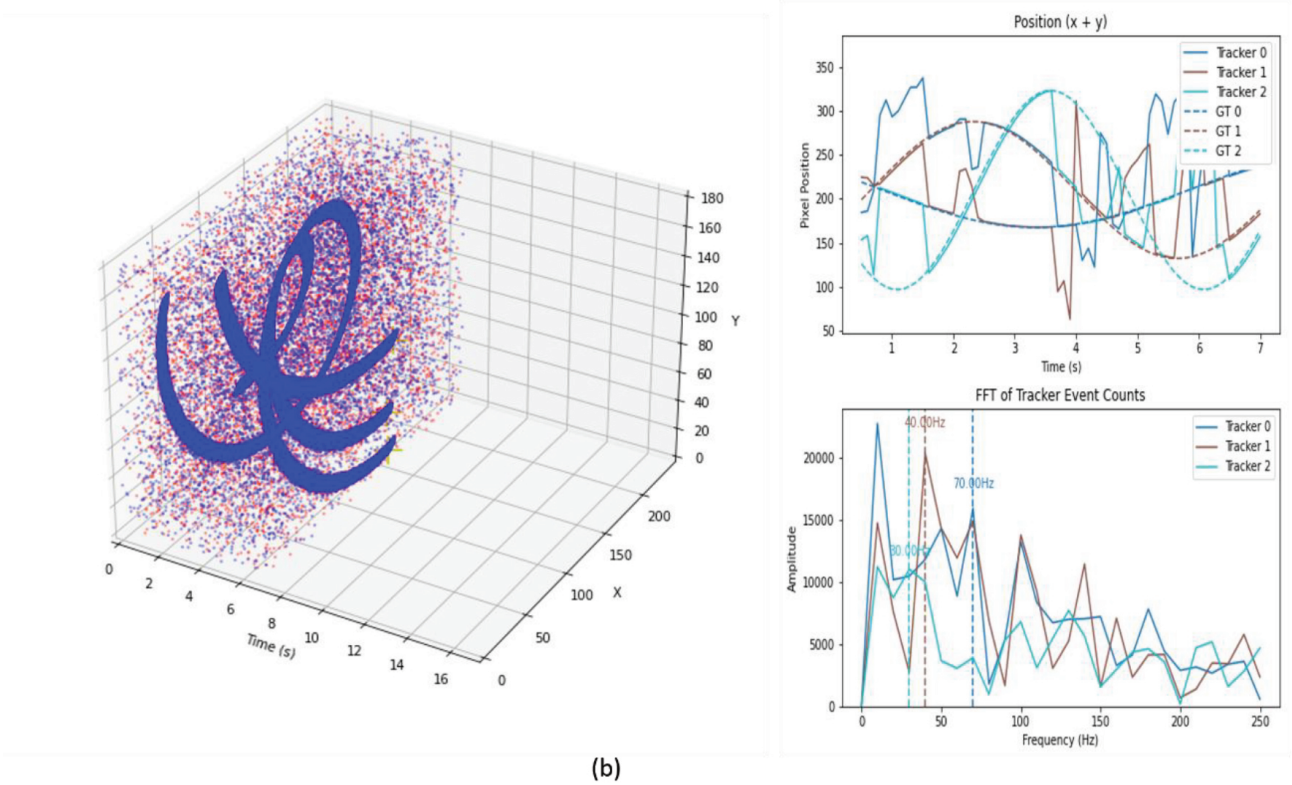


Fig. 3 3D visualizations of four simulated scenarios with three oscillating objects. (a) Stationary objects blinking at 1.0, 1.5, and 2.2 Hz; (b) circular motion with fast blinking at 30, 40, and 70 Hz; (c) moderate blink rates of 2.0, 4.5, and 7.0 Hz; (d) slower oscillations at 1.25, 2.0, and 3.5 Hz.



(a)



(b)

Fig. 4 *Continued*

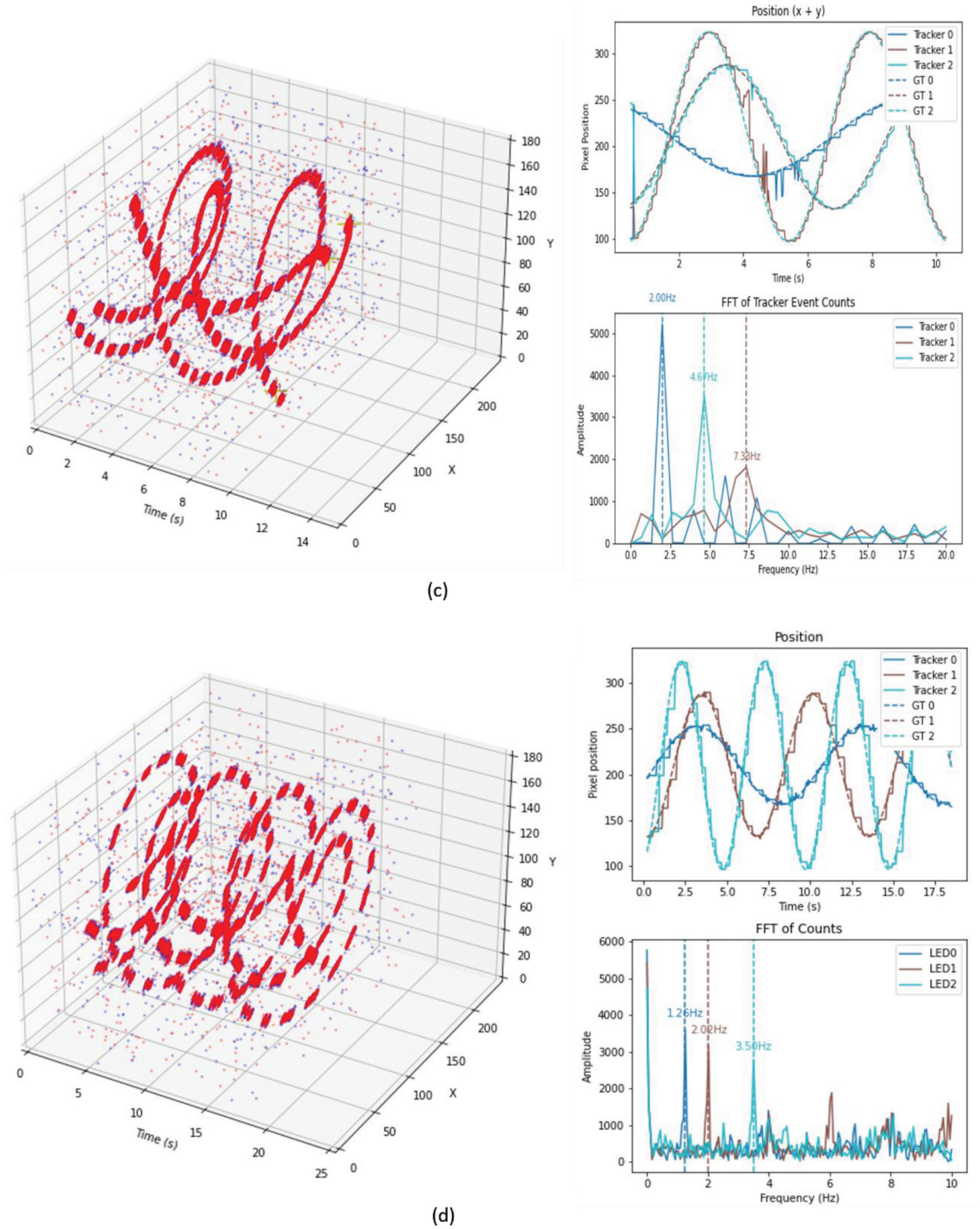


Fig. 4 Frequency estimation and spatial localization results for four scenarios: (a) static blinkers at 1.0, 1.5, 2.2 Hz (∼0.02 Hz error); (b) rapid motion at 30, 40, 70 Hz (∼0.5 Hz error with brief instabilities); (c) moderate motion at 2.0, 4.5, 7.0 Hz (∼0.03 Hz error); (d) slow oscillators at 1.25, 2.0, 3.5 Hz (∼0.015 Hz error).

Analysis

As shown in Figure 4, our PCA-based tracking filter accurately recovers the dominant oscillation frequencies in all four simulated scenarios. In the static case (Figure 5-a), three objects blinking at 1.0, 1.5, and 2.2 Hz yield a mean absolute frequency error of approximately 0.02 Hz ($\approx 1\%$ relative error). Under rapid motion, Figure 5-b (30, 40, and 70 Hz blinkers), the tracker experiences brief chaotic behavior when object trajectories intersect but stabilizes on its own, underscoring the framework's resilience. Despite these transient instabilities, frequency recovery remains largely unaffected (MAE ≈ 0.5 Hz, max error 1.1 Hz). We do observe that at higher frequencies the energy in harmonic peaks increases, which could occasionally lead to misidentifying a harmonic rather than the fundamental frequency. Moderate-motion scenarios (Figure 5-c at 2.0, 4.5, 7.0 Hz and Figure 5-d at 1.25, 2.0, 3.5 Hz) maintain high spectral fidelity, with MAE around 0.03 Hz and 0.015 Hz, respectively. In per-buffer mode, centroid localization remains robust, with an average spatial error of 1.2 px ($\sigma = 0.3$ px) across all scenarios. When Poisson-distributed background events are injected at 20% of the scene rate, both spectral and spatial errors increase by only about 10%, demonstrating the filter's robustness to realistic noise levels.

Importantly, the filter exhibits strong resilience to realistic sensor noise and partial overlap among event clouds. When Poisson-distributed background events are injected at 20% of the scene rate, both spatial and spectral metrics degrade by only about 10%, underscoring robustness to moderate noise levels. Together, these results demonstrate that our unified approach not only matches or exceeds prior single-object PCA methods in accuracy but also extends naturally to multiple, concurrent targets without sacrificing speed. By fusing histogram-based localization with sliding-window PCA updates, the tracker delivers high-precision spatial and spectral estimates in real-time, paving the way for rich, event-driven motion analyses in low-latency vision systems.

Laboratory Tests

To evaluate the tracker under real sensing conditions, we conducted a laboratory test using a single LED blinking at 1 Hz with a 50% duty cycle. The LED was fixed in front of the event camera and driven by a function generator. The LED was mounted in a fixed position directly in front of the event camera, providing a stable, high-contrast point source. This offers a clean, well-controlled source for validating frequency estimation.

The PCA-based tracker was implemented in a manner where incoming events were projected along the principal axis and accumulated in the histogram to refine the centroid location, as mentioned earlier. Figure 5 shows the histogram segment corresponding to approximately 4 seconds of recording, and it shows a clear, narrow peak at the LED position, demonstrating the method's ability to isolate a static point source under continuous blinking. An event snapshot created at the same time, shown in Figure 6, confirms this localization, with the detected centroid marked by a black cross.

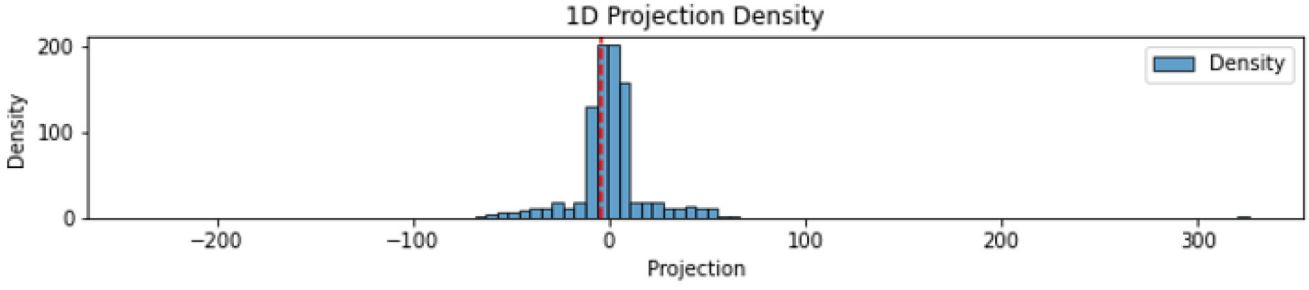


Fig. 5 Histogram of projected events over a 4-second recording of a static 1 Hz LED. A sharp, narrow peak indicates the detected LED position, demonstrating stable localization under periodic blinking.

We also examined the LED's detected pixel position over time (Figure 6). Although the source itself was fixed, small oscillations are visible in the estimated position trace. These appear as regular spikes and are attributed to the LED's on-off transitions. As the LED switches between on and off, subtle imbalances in event polarity can cause the centroid estimate to jitter slightly around its true position. Despite this artifact, the tracker remains stable. Finally, the frequency spectrum of the tracker's event counts reveals a dominant peak at 1 Hz, aligned with the programmed blink rate. These results confirm that the PCA tracker can not only localize a simple static source but also recover its periodic dynamics with high accuracy in real-time.

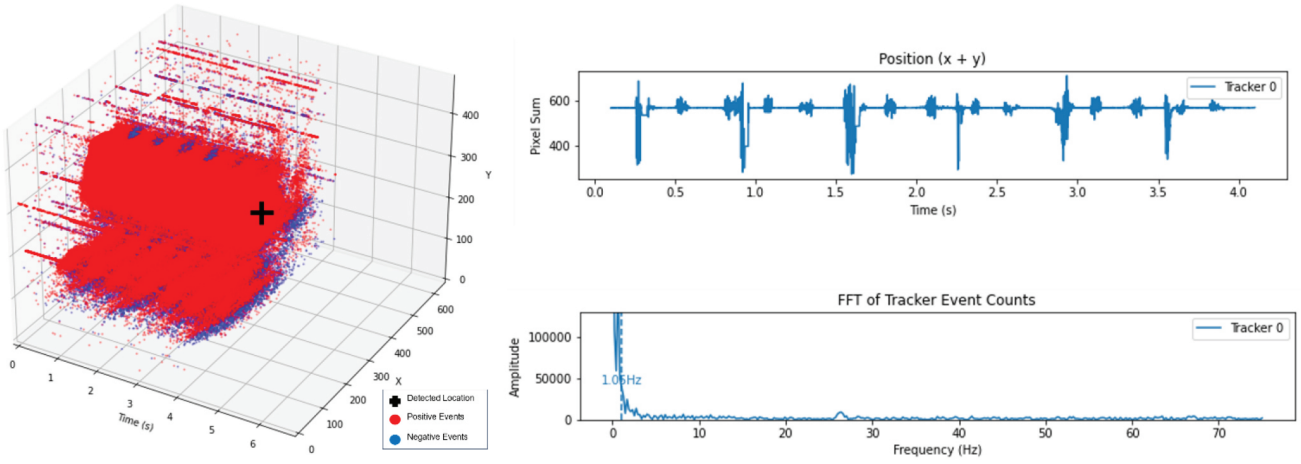


Fig. 6 Event snapshot and spectral analysis of the static 1 Hz LED. Left: detected centroid marked by a black cross on the event display. Right-top: estimated pixel position versus time, showing small oscillations linked to the LED's on-off transitions. Right-bottom: frequency spectrum of tracker event counts with a clear dominant peak at 1 Hz.

Conclusions

This paper presented a PCA-based framework for object tracking and frequency identification in event-driven sensing systems. In particular, the method was assessed through simulation studies and a laboratory experiment involving oscillating light sources, where it was shown to provide accurate estimates of spatial trajectories and dominant motion frequencies. The framework relied on continuous updates of principal directions combined with histogram refinement and did not require reconstructed frames or trained learning models. It was further observed that the approach handled overlapping motion paths and noise without significant degradation in performance. These findings indicate that PCA-based processing constitutes a reliable basis for low latency motion analysis in embedded applications.

Acknowledgments This paper was supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project number 20250620MFR, as well as Innovation Hub Sprint funding sources internal to Los Alamos National Laboratory. Los Alamos National Laboratory is operated by Triad National Security, LLC, for the National Nuclear Security Administration of US Department of Energy (Contract No. 89233218CNA000001). This work was also supported by the Department of Energy – Environmental Management Office with R&D Award: 277483. The DVS/DAVIS technology was developed by the Sensors group of the Institute of Neuroinformatics (University of Zurich and ETH Zurich), which was funded by the EU FP7 SeeBetter project (grant 270324).

References

1. G. Gallego, T. Delbrück, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. J. Davison, J. Conradt, K. Daniilidis, and D. Scaramuzza, "Event-based Vision: A Survey," IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, no. 1, pp. 154–180 (2020)
2. Luming Wang, Hao Shi, Xiaoting Yin, Kailun Yang, Kaiwei Wang, and Jian Bai, "EgoEvGesture: Gesture Recognition Based on Egocentric Event Camera," CoRR, abs/2503.12419v2 (2025)
3. X. Zheng, Y. Liu, Y. Lu, T. Hua, T. Pan, W. Zhang, D. Tao, and L. Wang, "Deep Learning for Event-based Vision: A Comprehensive Survey and Benchmarks," CoRR, abs/2302.08890 (2023)
4. Mahmoud Z. Khairallah, Fabien Bonardi, David Roussel, and Samia Bouchafa, "PCA Event-Based Optical Flow for Visual Odometry," CoRR, 2021
5. B. Ramesh, A. Ussa, L. Della Vedova, H. Yang, and G. Orchard, "PCA-RECT: An Energy-efficient Object Detection Approach for Event Cameras," CoRR, abs/1904.12665v1
6. F. Hamann, D. Gehrig, F. Febryanto, K. Daniilidis, and G. Gallego, "ETAP: Event-based Tracking of Any Point," in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 2025
7. A. Sabater, L. Montesano, and A. C. Murillo, "Event Transformer+: A multi-purpose solution for efficient event data processing," CoRR, abs/2211.12222v2
8. B. Pfrommer, "Frequency Cam: Imaging Periodic Signals in Real-Time," CoRR, abs/2211.00198

9. C. J. Benson and M. Holzinger, “Simulation and Analysis of Event Camera Data for Non-resolved Objects,” Journal of the Astronautical Sciences, vol. 71, article 3
10. T. Fischer and M. Milford, “Event-based visual place recognition with ensembles of temporal windows,” IEEE Robotics and Automation Letters, vol. 5, no. 4, pp. 6924–6931 (2020)
11. I. T. Jolliffe, “Principal Component Analysis,” 2nd ed., Springer (2002)
12. M. Brand, “Fast Low-Rank Modifications of the Thin Singular Value Decomposition,” MERL Tech. Rep. TR2006-059 (2006)

Chapter 7
3D Motion Detection Using Computer Vision During Earthquakes

Andrés Garreaud, Ruben Boroschek, and Domingo Mery

Abstract In recent years, the field of Computer Vision has gained significant traction in civil engineering, largely due to advancements in techniques and the increased availability of data from cameras and unmanned aerial vehicles (UAVs). This work presents a series of controlled experiments in which objects are recorded on video while following predefined motion patterns on a six-degree-of-freedom shake table. The dimensions of the objects are fully known, allowing a single camera to be used for tracking. This case is important because understanding the benefits and limitations of single-camera video during seismic events is highly relevant. Handheld cameras and surveillance systems inside buildings generally lack binocular (stereo) vision. The motion patterns range from simple movements to complex ones derived from actual seismic recordings, and the amplitudes vary from low-level ambient vibrations to severe earthquake motions. Using Computer Vision techniques, motion is first estimated in 2D space through feature tracking across video frames, then transformed into 3D space using a calibrated camera model. Calibration is performed using the known object dimensions. We assess the accuracy of the proposed approach by comparing the estimated displacements with the predefined motion patterns. The limitations of the method are identified and discussed. These results contribute to the development of more robust and precise Computer Vision workflows, which can be applied to more complex environments, such as real-world seismic structural monitoring.

Keywords Displacement estimation · 3D motion detection · single-camera tracking · computer vision · seismic monitoring

Introduction

Over the past few years, computer vision techniques have become increasingly common for monitoring and inspecting structures [1, 2, 3]. Among the most relevant applications is the estimation of structural displacements, traditionally addressed through 2D methods such as homography or the scale factor approach [4, 5]. These techniques, however, are limited to planar motion. More complex challenges arise when estimating 3D displacements, typically requiring two or more cameras for triangulation [6].

In contrast, estimating 3D motion with a single camera is far more challenging due to the lack of stereoscopic information. Recent studies have explored neural networks [7], segmentation [8, 9], and calibration models to address this issue [10]. Building on this line, the present work proposes a method that combines monocular vision with optimization techniques, integrating prior knowledge of object geometry to estimate 3D displacements. The method is experimentally validated using a six-degree-of-freedom shake table, where a rigid object with known dimensions is subjected to predefined motion patterns and recorded with fixed cameras. AprilTags were included in the setup but were not used for this extended Abstract.

Andrés Garreaud
Graduate Student Department of Civil Engineering University of Chile Av. Blanco Encalada 2002, Santiago, Chile
e-mail: garreaudandres@gmail.com

Ruben Boroschek
Professor Department of Civil Engineering University of Chile, Av. Blanco Encalada 2002, Santiago, Chile
e-mail: rboroschek@uchile.cl

Domingo Mery
Professor Department of Computer Science Pontificia Universidad Catolica of Chile, Av. Vicuña Mackenna 4860, Macul, Santiago, Chile
e-mail: domingo.mery@uc.cl

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

The proposed method begins with the detection and tracking of features on the studied object. Several detection algorithms were evaluated (FAST [11], SURF [12], and Harris Corner [13]); the Harris detector was selected in this case. Feature tracking across frames was then performed using the KLT algorithm [14], yielding the 2D trajectories of all points. The relationship between a 3D point with coordinates (X, Y, Z) and its 2D projection was implemented using a novel methodology that optimally finds the intrinsic and extrinsic parameters for this case. This methodology is presented in [15].

Laboratory Validation

A six-degree-of-freedom shaker table is employed to conduct a series of experiments aimed at verifying and validating the proposed method. A black solid box and two 3D-printed cubes were placed on the table, and all objects were rigidly fixed to ensure that they moved together, strictly following the imposed motion patterns. The equipment is shown in Figure 1, while their arrangement is depicted in Figure 2.

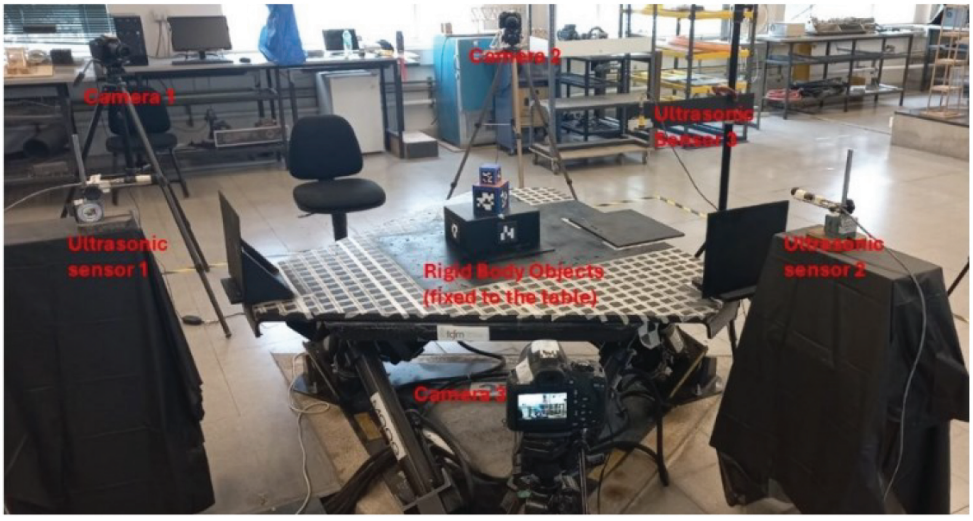


Fig. 1 Laboratory experimental setup

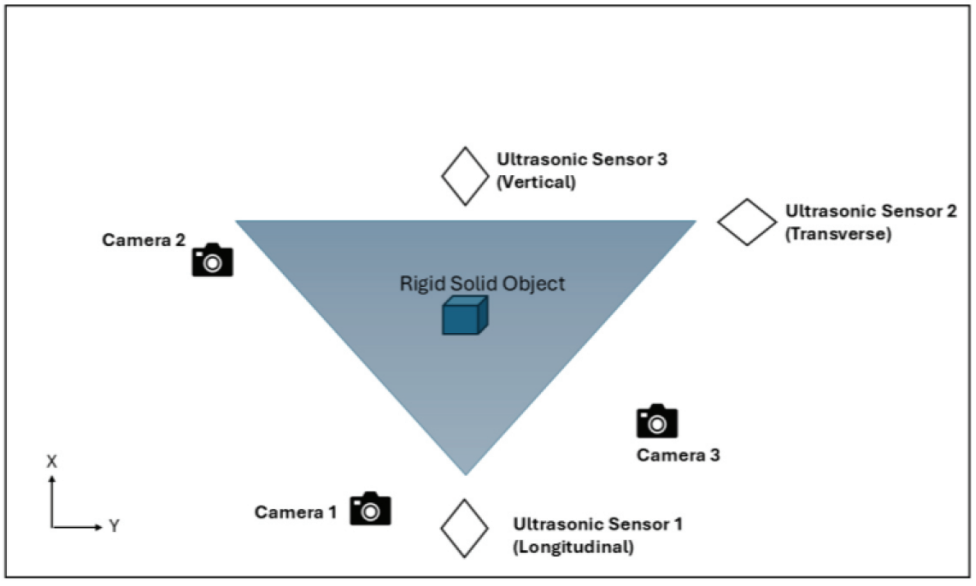


Fig. 2 Schematic representation of the experimental setup

Three Canon Rebel T6 cameras were mounted on tripods to record different faces of the object, operating at 1920×1080p and 30 fps with manual settings. One camera at a time was used to test the methodology. Table motion was independently measured using three ultrasonic sensors (BANNER S18U, 200 Hz), one per axis. Regarding the motion patterns, two records were employed. The first was manually generated and consisted of three displacement pulses in the principal directions (X, Y, and Z), each lasting 5 seconds with a maximum amplitude of 5 cm. Similar results were obtain from the different tests. The second record corresponds to seismic data from the 2010 Chile earthquake. Only the results for the earthquake-induced 3D motions are presented.

Results

For the 2010 Chile Earthquake Constitución record, features were independently tracked by Cameras 1–3. The estimated displacements (Figures 3–5) closely follow the sensor data, reproducing both strong-motion peaks and slower phases with high accuracy. Cameras 1 and 2 show excellent agreement in all directions, while Camera 3 provides precise results in X and Z but underestimates amplitudes along Y during the strong-motion phase. Quantitative results (Table 1) indicate RMSE values below 1 cm (≈1% error), with maximum deviations of 2.06% and 3.69% in the X and Y directions of Camera 3, respectively. All CCF values remain above 0.95, confirming the reliability of the method.

Table 1 Displacement measurement errors for seismic recording test

		Camera1			Camera 2			Camera3	
	X	Y	Z	X	Y	Z	X	Y	Z
RMSE (cm)	0.09	0.19	0.11	0.13	0.15	0.07	0.23	0.65	0.05
NRMSE (%)	0.9	1.1	1.9	1.2	0.87	1.3	2.06	3.69	0.92
CCF (-)	0.99	0.99	0.99	0.99	0.99	0.99	0.98	0.95	0.99

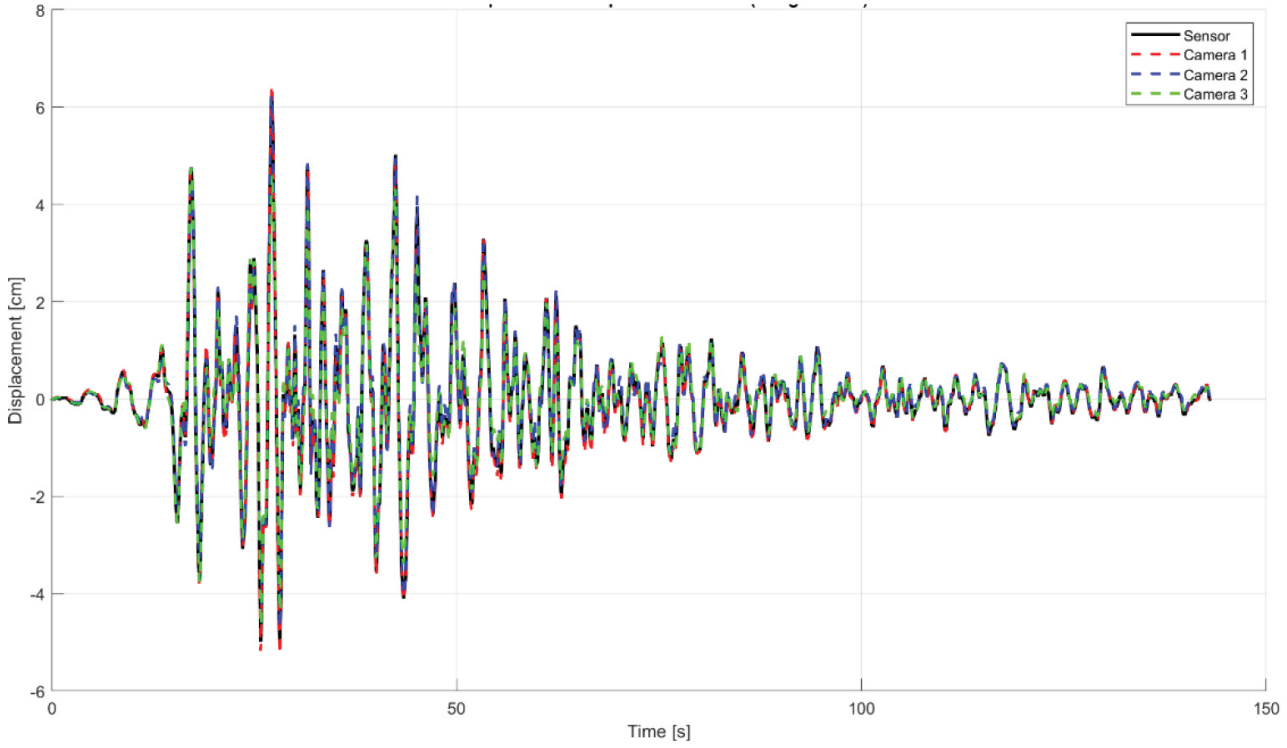


Fig. 3 Comparison of displacement results along the X direction

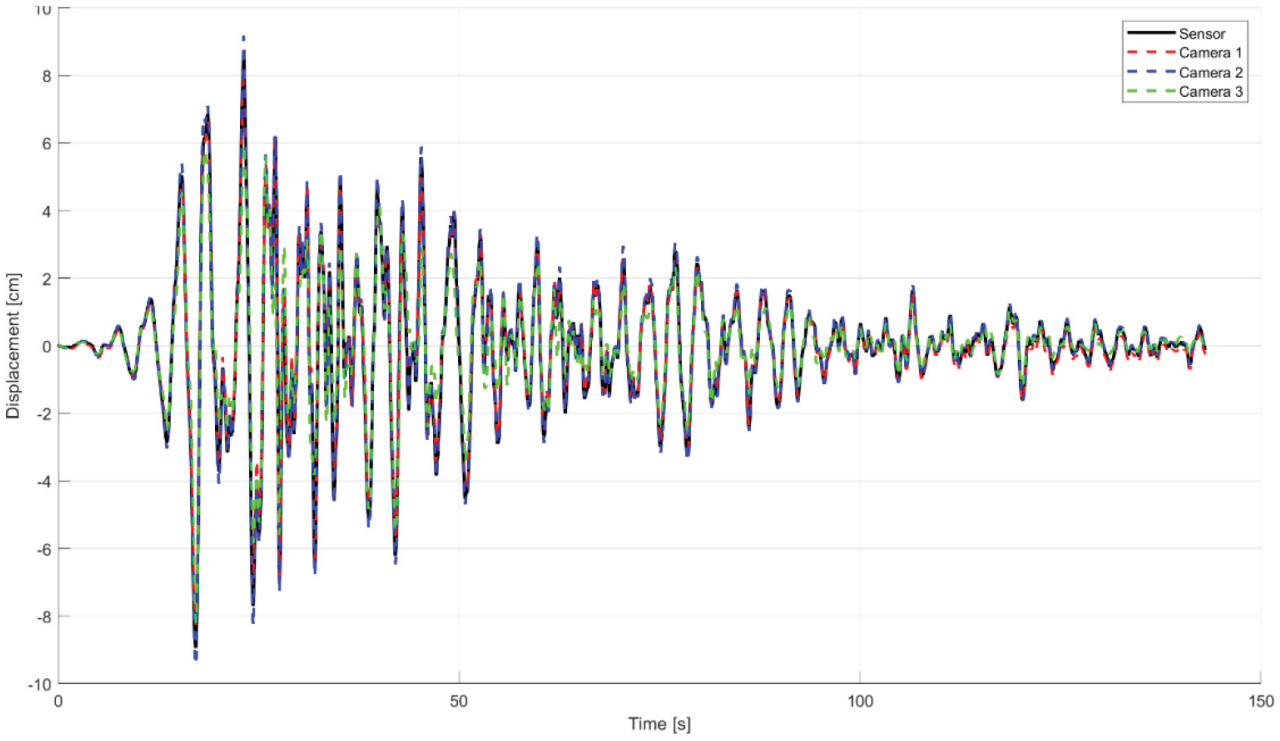


Fig. 4 Comparison of displacement results along the Y direction

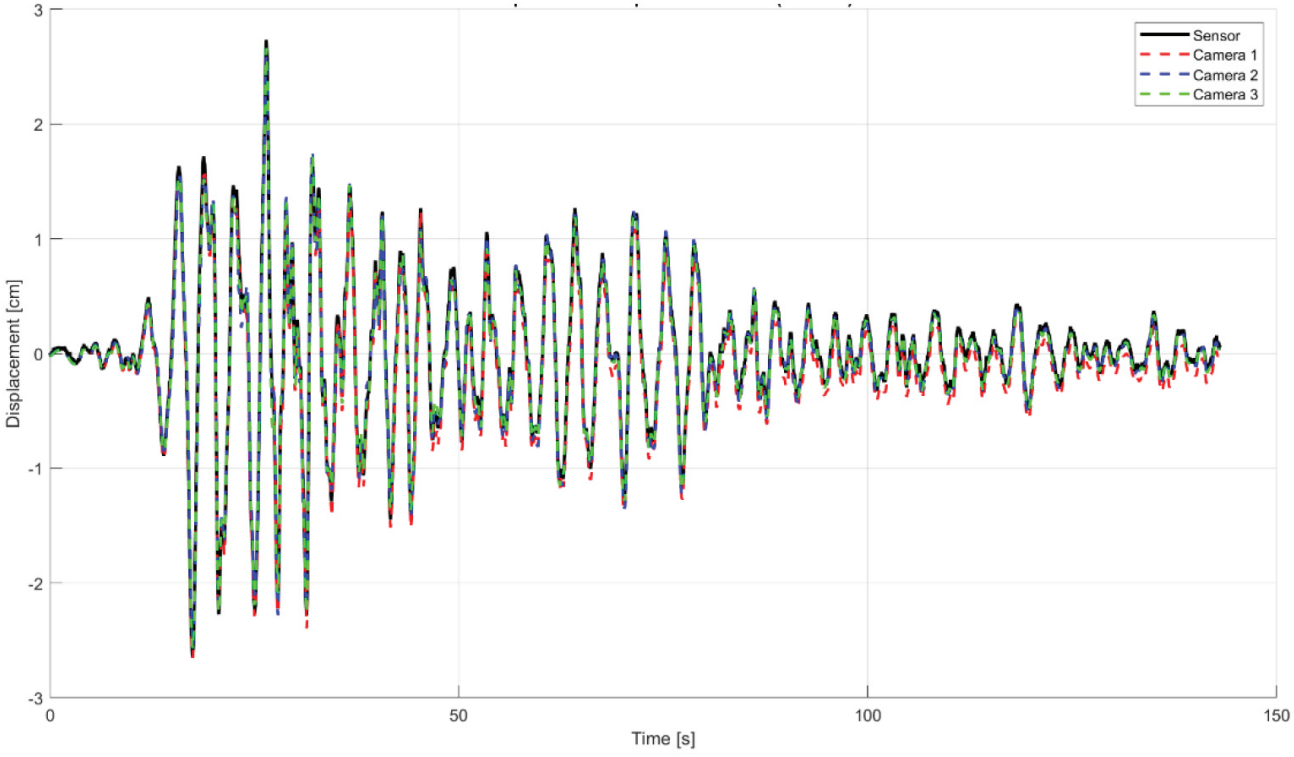


Fig. 5 Comparison of displacement results along the Z direction

Conclusion

This study demonstrates that single cameras can be effectively used to track known objects during severe earthquakes with high accuracy. The proposed method, based on monocular vision and optimization techniques, successfully reproduces 3D displacements consistent with sensor measurements, even under strong-motion conditions. We are applying this method to real videos during earthquakes, and the results will be presented in a forthcoming working paper.

References

1. Spencer, B., Hoskere, V., and Narazaki, Y., "Advances in computer vision-based civil infrastructure inspection and monitoring," *Engineering*, vol. 5, no. 2, pp. 199-222, 2019.
2. Xu, Y., and Brownjohn, J.M., "Review of machine-vision based methodologies for displacement measurement in civil structures," *Journal of Civil Structural Health Monitoring*, vol. 8, no. 1, pp. 91-110, 2018.
3. Baqersad, J., Poozesh, P., Niezrecki, C., and Avitabile, P., "Photogrammetry and optical methods in structural dynamics-a review," *Mechanical Systems and Signal Processing*, vol. 86, pp. 17-34, 2017.
4. Lee, J., Cho, S., and Sim, S., "Monocular vision-based displacement measurement system robust to angle and distance using homography," in *6th international conference on advances experimental structural engineering*, 2015.
5. Feng, D., Feng, M.Q., Ozer, E., and Fukuda, Y., " A vision-based sensor for noncontact structural displacement measurement," *Sensors*, vol. 15, no. 7, pp. 16557-16575, 2015.
6. Feudo, S., Han, Y., Cumunel, G., Hoult, R., Bertholet, A., Garnier, D., Candeias, P., Correia, A., and Pacheco de Almeida, J., "Video tracking of targets for vibration measurement of large-scale structures under seismic excitation," *Earthquake Engineering Structural Dynamics*, vol. 54, pp. 2106-2120, 2025.
7. Chujin, S., Donglian, G., and Xinzheng, L., "Three-dimensional structural displacement measurement using monocular vision and deep learning based pose estimation," *Mechanical Systems and Signal Processing*, vol. 190, 2023.
8. Li, Q., Shao, Y., Li, L., Li, J., and Hao, H., "Advancements in 3d displacement measurement for civil structures: A monocular vision approach with moving cameras," *Measurement*, vol. 242, 2025.
9. Shao, Y., Li, L., Li, J., Li, Q., An, S., and Hao, H., "3d displacement measurement using a single-camera and mesh deformation neural network," *Engineering Structures*, vol. 318, 2024.
10. Narazaki, Y., Gomez, F., Hoskere, V., Smith, M., and Spencer, B., "Efficient development of vision-based dense three-dimensional displacement measurement algorithms using physics-based graphics models," *Structural Health Monitoring*, vol. 20, no. 4, pp. 1841-1863, 2020.
11. Trajkovic, M., and Hedley, M., "Fast corner detection," Image and Vision Computing, vol. 16, no. 2, pp. 75-87, 1998.
12. Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L., "Speeded-up robust features (surf)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346-359, 2008.
13. Harris, C., and Stephens, M., "A combined corner and edge detector," in *Proceedings of the fourth Alvey vision conference*, Manchester, UK, 1988.
14. Lucas, B., and Kanade, T., "An iterative image registration technique with an application to stereo vision," in *7th international joint conference on Artificial*, Vancouver, Canada, 1981.
15. Garreaud, A., Boroschek, R., and Mery, D., "Detection of 3D Motion Using Computer Vision During Earthquakes: Validation and Application with Single Cameras," *Paper in preparation*, 2025.

Chapter 8
Vibration Measurement of a Rotating Structure Using an Image-Based Tracking Continuously Scanning Laser Doppler Vibrometer with a Misalignment Angle

L.F. Lyu, W.D. Zhu, D. T. Will, and K. Yuan

Abstract A novel misalignment angle correction method is developed for an image-based tracking continuously scanning laser Doppler vibrometer (CSLDV) for tracking and scanning a rotating structure with a large misalignment angle between its rotation axis and the optical axis of the tracking CSLDV's camera. In this method, the rotation trajectory of the structure is an ellipse, where the position of its center, the lengths of its semi-major axis and semi-minor axis, and the angular position of its major axis are used to determine the misalignment angle and the endpoints of the scan paths on the rotating structure. In the experimental investigation, an aluminum blade is attached to a hub on a frame, where the frame is turned to misalignment angles of 0°, 10°, 20°, and 30° between the tracking CSLDV and the rotating blade. The blade rotates at a constant speed driven by a motor inside of the hub, and the blade is excited by air flow from a fan placed behind the rotating structure. The rotating blade with different misalignment angles is tracked and scanned by the tracking CSLDV and the measured responses are processed by the improved demodulation method to estimate the modal parameters, including the damped natural frequencies and undamped mode shapes. The estimated rotation speeds and modal parameters for each angular misalignment are compared with each other, as well as to the results when angular misalignment is not considered, where the novel method is shown to produce better measurement results.

Keywords rotation tracking · misalignment angle correction

Introduction

A laser Doppler vibrometer (LDV) is a well-developed and commercialized technology for non-contact vibration measurement. The LDV can accurately measure the surface velocity of a point on a structure along the direction of its laser beam, which can then be used as the vibration response at that point [1]. To measure the responses of the structure on a measurement grid, researchers may need to use multiple LDVs or move the LDV to multiple positions, which can be expensive and time consuming. A scanning laser Doppler vibrometer (SLDV) was developed to provide measurements on a measurement grid [1–3]. In an SLDV, a scanner with a set of X- and Y-mirrors is installed in front of the LDV to reflect its laser beam. The rotation axes of the two mirrors in the scanner are orthogonal, and the rotation angles of the mirrors are considered linearly related to the horizontal and vertical position coordinates of the laser spot on the structure. The location of the laser spot can be controlled by rotating the mirrors in the scanner. The SLDV measures the responses on a measurement grid in a step scanning manner, i.e., its laser spot stays at a fixed point for a period of time and then moves to the next point [4, 5]. To increase the efficiency in measuring the vibration responses on a dense measurement grid, a continuously scanning laser Doppler vibrometer (CSLDV) was developed [6–9]. A continuous trajectory is generated on a structure and the laser spot is swept along its surface by continuously rotating the mirrors of the scanner in the CSLDV. Thus, the CSLDV can estimate mode shapes and operational deflection shapes (ODSs) with high spatial resolution. Further, to measure the three-dimensional (3D) vibration of a structure, a 3D SLDV was developed and extended to a 3D CSLDV by combining three individual CSLDVs [10–14].

L.F. Lyu · W.D. Zhu(✉) · D. T. Will · K. Yuan
Department of Mechanical Engineering University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250
e-mail: linfenl1@umbc.edu; wzhu@umbc.edu; dwill1@umbc.edu; kyuan1@umbc.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

The CSLDV is used to measure the responses of a stationary structure under different excitations, and different operational modal analysis (OMA) methods were developed to process these measured responses. A demodulation method and a polynomial method were developed to process the measurements of a CSLDV and estimate the ODSs of a structure under sinusoidal excitation [15, 16], and the estimated ODSs were used for baseline-free damage detection [17–21]. The demodulation method was improved to allow for processing of the measured response of structures under random excitation [11, 22, 23]. A lifting method was developed to estimate the undamped modes shapes of a structure under random excitation [24–26].

To track and scan a rotating structure without attaching an encoder or mark, an edge detection image processing method with a distance condition was developed for an image-based tracking CSLDV to process the images of the rotating structure and determine its real-time position [27]. The neighboring frame subtraction technique [28] and the background subtraction technique [29] image processing algorithms were developed to improve the demodulation method so that a rotating structure with a complex background can be tracked. These image-based tracking methods do not require the tracking CSLDV to be aligned with the rotation axis of the structure being tracked. However, the estimated modal parameters of a rotating structure with a large misalignment angle between the rotation axis and the image-based tracking CSLDV can have some differences when compared to the modal parameters of a rotating structure with no misalignment angle [28]. To reduce the effects of a large misalignment angle on the estimated modal parameters of a rotating structure using the image-based tracking CSLDV, a misalignment angle correction method is needed.

The remainder of this work is as follows. A misalignment angle correction method for an image with a large misalignment angle is introduced in Sec. 2.1, a 1D scan scheme based on the misalignment angle correction method is introduced in Sec. 2.2, and a signal processing method for analyzing the measured responses is introduced in Sec. 2.3. An experimental setup for tracking and scanning a rotating blade with different misalignment angles is introduced in Sec. 3.1, misalignment angle correction results are presented in Sec. 3.2, OMA results of the rotating blade with different misalignment angles are presented in Sec. 0, and results of using the proposed method on a rotating blade in an outdoor environment are presented in Sec. 3.4. Some conclusions are presented in Sec. 4.

Methodology

Misalignment Angle Correction Method for an Image with a Large Misalignment Angle

The boundary of the trajectory of a rotating structure in its rotation plane is a circle whose radius is the radius of the rotating structure. If the boundary that can be considered as a circular object turns about a line in its radial direction, its parallel projection on the plane of the circle is an ellipse (Figure 1), where the lengths of its semi-major axis and semi-minor axis are a and b, respectively. Figure 1 depicts the top and front views of both the original circular object and the turned circular object. The turning angle of the circular object can be calculated by

$$\theta = \arccos(b/a) \tag{1}$$

In this research, the boundary of the trajectory of the rotating blades, whose hub is mounted on an aluminum frame, is considered as a circular object (Figure 2). The trajectory of the rotating blades shown in Figure 2(b) is an overlay of 500 images of the rotating blades captured by the camera in an image-based tracking CSLDV with a constant frame rate of 25 frames per second. When the rotation axis of the rotating blades is aligned with the optical axis of the camera, the boundary of the trajectory is a circle and Eq. (1) can be used to find that $\theta = 0°$.

When the rotation axis of the rotating blades is not aligned with the optical axis of the camera, the boundary of the trajectory of the rotating blades in the captured images is considered as an ellipse (Figure 3), where the frame supporting the rotating blades is turned 30°. The parameters of the ellipse shown in Figure 3 are determined by the ellipse detection algorithm in Ref. [30], where the pixel coordinates of the ellipse center are (571.26, 508.62), the lengths of the semi-major axis and semi-minor axis are 455.99 pixels and 399.16 pixels, respectively, and the angle between the semi-major axis and the X-axis is −88.47°. Using Eq. (1), the calculated angle between the rotation axis of the blades and the optical axis of the camera is 28.91°.

The parameters of the ellipse can be used to generate an accurate scan path on a rotating blade with a large misalignment angle of 30°. To generate a scan path on the rotating blade, the first step is to determine the position of the rotating blade in a captured image from the camera. In this work, the distance condition in Ref. [28] and the background averaging technique in Ref. [29] are used to determine the position of an edge point of the rotating blade. The details of the image processing method for tracking the rotating blade can be found in Ref. [27–29].

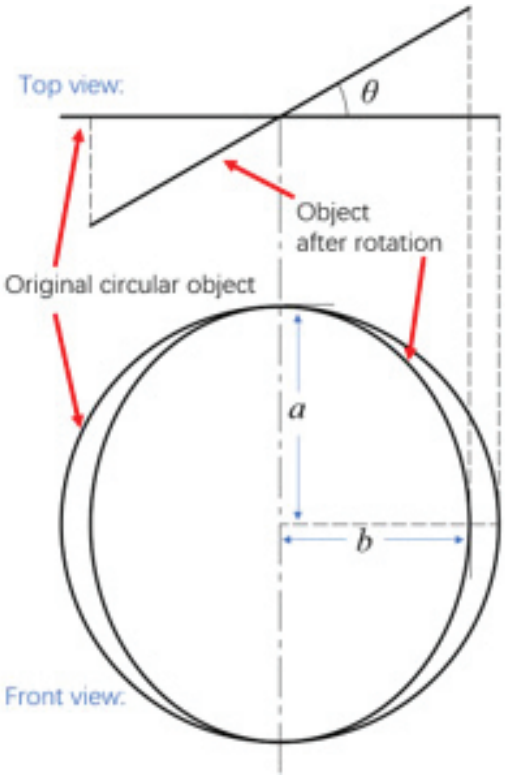


Fig. 1 Schematic of the parallel projection of the rotated circular object.



(a) (b)

Fig. 2 (a) Image of the rotating blades, and (b) the image of the trajectory of the rotating blades in (a).

In Figure 4, the middle line of the rotating blade is considered as a line passing through its rotation center, and the angle between the middle line and the X-axis is determined by the position of the edge point and the angle α. The equation of the middle line is given by

$$(y-n) = (x-m)\tan\beta \tag{2}$$

where (m, n) are the coordinates of the rotation center, and β is the angle between the middle line and the X-axis. The equation of the ellipse is given by

$$\frac{[(x-p)\cos\varphi+(y-q)\sin\varphi]^2}{a^2} + \frac{[(p-x)\sin\varphi+(y-q)\cos\varphi]^2}{b^2} = 1 \tag{3}$$

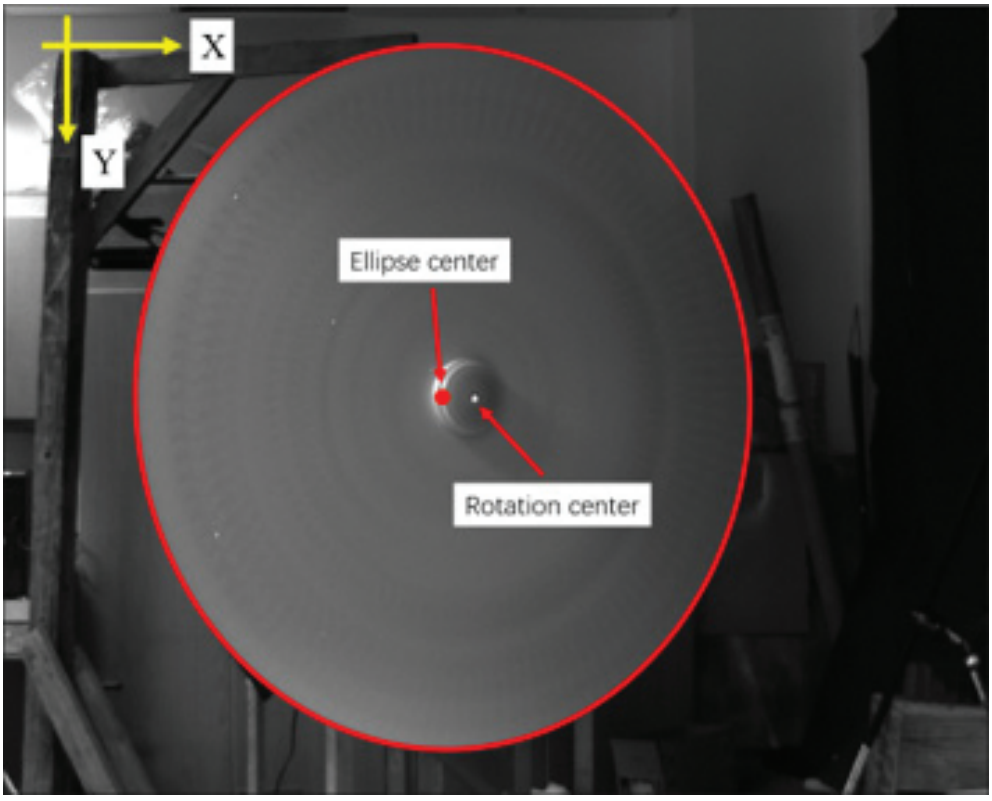


Fig. 3 Image of the trajectory of the rotating blades with its estimated boundary.

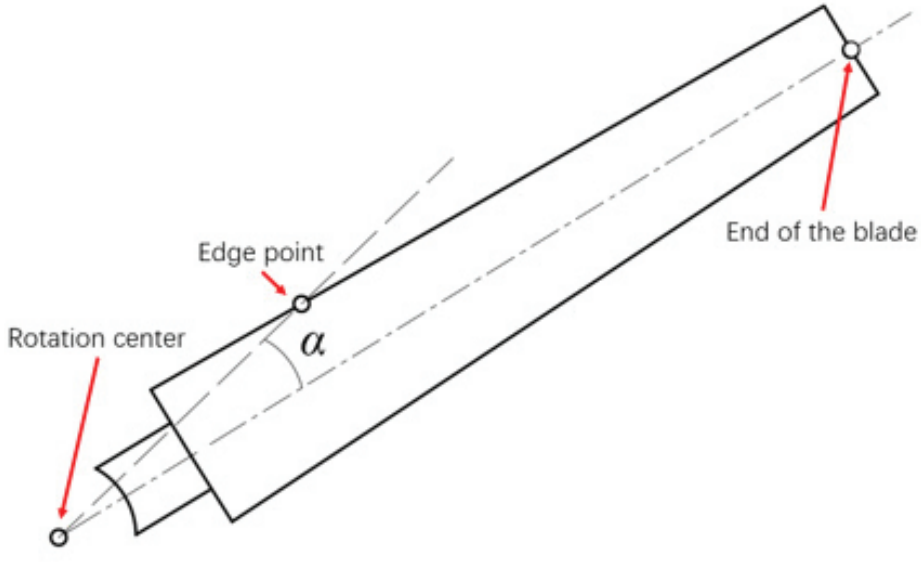


Fig. 4 Schematic of the rotating blade.

where (p, q) are the coordinates of the ellipse center, and φ is the angle between the major axis of the ellipse and the X-axis (Figure 5). The intersection point of the ellipse and the middle line is the end of the rotating blade, and its position is used to generate accurate scan paths on the rotating blade. Eq. (2) can be rewritten as

$$\begin{cases} x = m + t\cos\beta \\ y = n + t\sin\beta \end{cases} \tag{4}$$

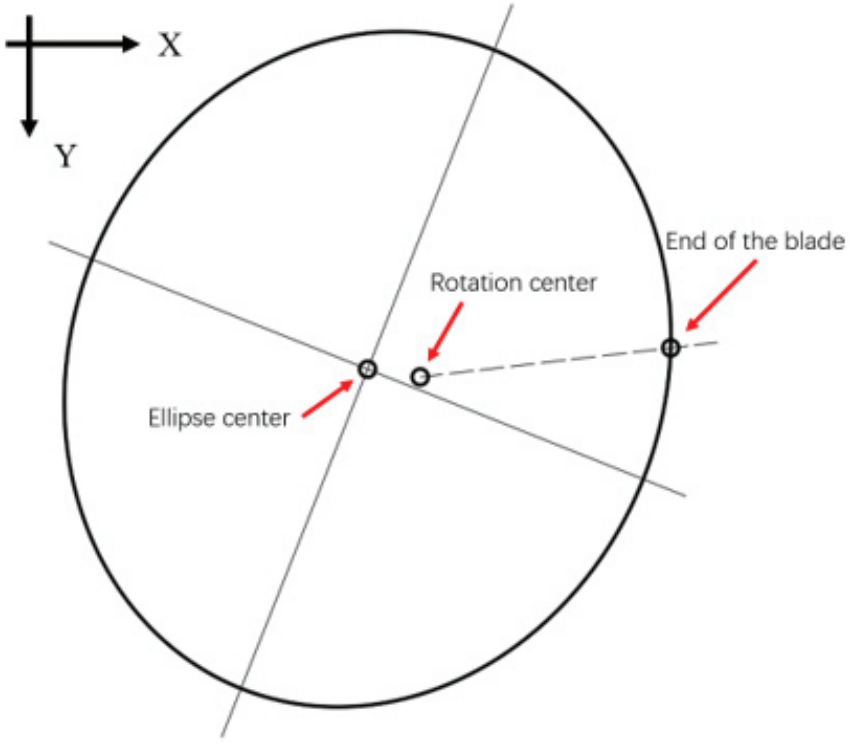


Fig. 5 Schematic of the trajectory boundary of the rotating blade with a large misalignment angle.

where t is a positive parameter. Substituting Eq. (4) into Eq. (3) yields

$$\frac{\left[(m+t\cos\beta-p)\cos\varphi+(n+t\sin\beta-q)\sin\varphi\right]^2}{a^2}+\frac{\left[(p-m-t\cos\beta)\sin\varphi+(n+t\sin\beta-q)\cos\varphi\right]^2}{b^2}=1 \tag{5}$$

Eq. (5) can be derived as

$$At^2+Bt+C=0 \tag{6}$$

where $A=\frac{\cos^2(\beta-\varphi)}{a^2}+\frac{\sin^2(\beta-\varphi)}{b^2}$, $B=\frac{2\cos(\beta-\varphi)[(m-p)\cos\varphi+(n-q)\sin\varphi]}{a^2}-\frac{2\sin(\beta-\varphi)[(p-m)\sin\varphi+(n-q)\cos\varphi]}{b^2}$, and $C=\frac{[(m-p)\cos\varphi+(n-q)\sin\varphi]^2}{a^2}+\frac{[(p-m)\sin\varphi+(n-q)\cos\varphi]^2}{b^2}-1$. The solutions to Eq. (6) are $t_1=\frac{-\mathrm{B}+\sqrt{B^2-4AC}}{2A}$ and $t_2=\frac{-\mathrm{B}-\sqrt{B^2-4AC}}{2A}$, where t_1 is chosen to be substituted into Eq. (4) since $t>0$. Once the position of the intersection point is obtained, it can be used to generate the endpoints of a scan path on the rotating structure.

1D Scan Scheme

Let the estimated position of the end of the rotating blade be (x_{end}, y_{end}). The position of an endpoint of the scan path on the rotating blade can be calculated by

$$\begin{cases} x=m+c(x_{\mathrm{end}}-m) \\ y=n+c(y_{\mathrm{end}}-n) \end{cases} \tag{7}$$

where c is a coefficient. By changing the value of c, the endpoint near the hub of the rotating blade and the endpoint near the free end of the rotating blade can be generated. The generated endpoints are shown in Figure 6(a), where 20 endpoints are generated near the boundary of the trajectory of the rotating blades and 20 endpoints are generated near the hub of the rotating blades. Using the misalignment angle correction method and considering the misalignment angle, the scan paths are generated based on the position coordinates of the endpoints and 100 scan paths are shown in Figure 6(b) as dashed lines. Conversely, if the legacy method is used where the misalignment angle is not considered when generating the scan paths [28], the scan paths may not align with the trajectory of the rotating blades (Figure 6(c)). The scan paths shown in Figure 6(c) are generated by the angular position of the middle line of the rotating blade only, and the boundary of the scan paths is a circle. As shown in Figure 6(c), portions of some scan paths are outside of the boundary of the trajectory of the rotating blades, which will add noise to the measurements of the image-based tracking CSLDV. Further, if the scan paths are

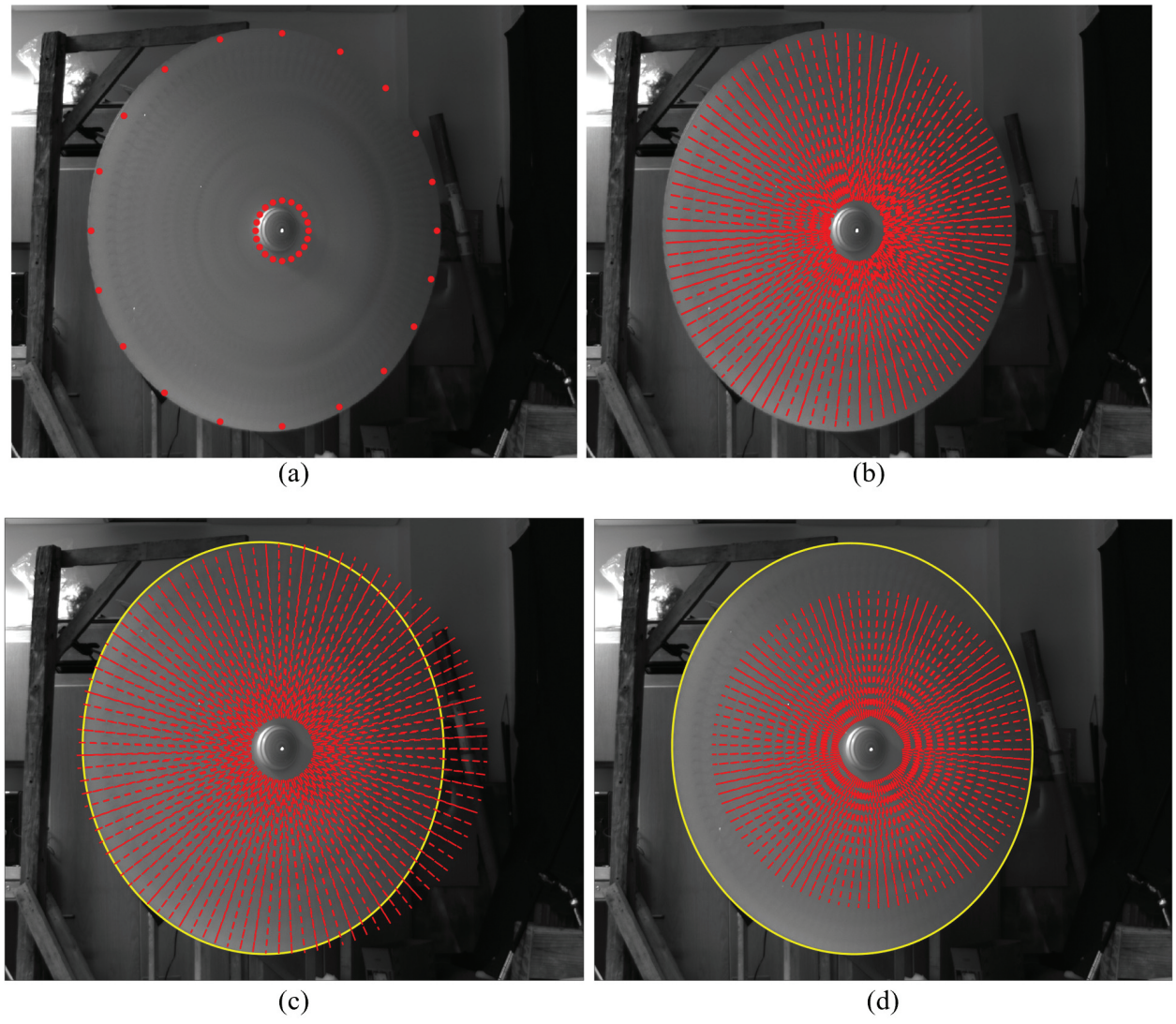


Fig. 6 (a) Endpoints of the scan paths generated on the trajectory of the rotating blades, (b) the scan paths generated on the trajectory of the rotating blades based on the estimated position coordinates of the end of the rotating blade, (c) the scan paths generated on the trajectory of the rotating blades based on the angular position of the middle line of the rotating blade, and (d) the shortened scan paths in (c).

shortened such that they are within the boundary of the trajectory of the rotating blades, the scan paths cannot fully capture the entire trajectory of the rotating blades (Fig. 6(d)), leading to inaccurate OMA results.

Signal Processing Methods for Image-Based Tracking CSLDV Measurements

The position of the shifted edge point on the middle line of the rotating blade can be used to calculate the rotation speed of the blade. To increase the accuracy of the estimated rotation speed, the middle line on the elliptical trajectory of the rotating blade can be transformed to the middle line on the original circular trajectory. The solid elliptical boundary line and the dashed circular boundary line are shown in Fig. 7(a), where the major axis of the ellipse is considered the radial direction at which the original circular boundary rotates about. In Fig. 7(b), point *O* is the ellipse center, points *A* and *B* are two estimated positions of the end of the rotating blade from two neighboring frames captured by the camera, and points *A'* and *B'* are points on the original circular boundary that correspond to points *A* and *B*, respectively. The line passing through points *A* and *A'* and the line passing through points *B* and *B'* are perpendicular to the major axis of the elliptical boundary. Let the angle between line *OA* and the X-axis be β_A, and the angle between line *OA* and the major axis of the ellipse is $\beta_A - \varphi$. The

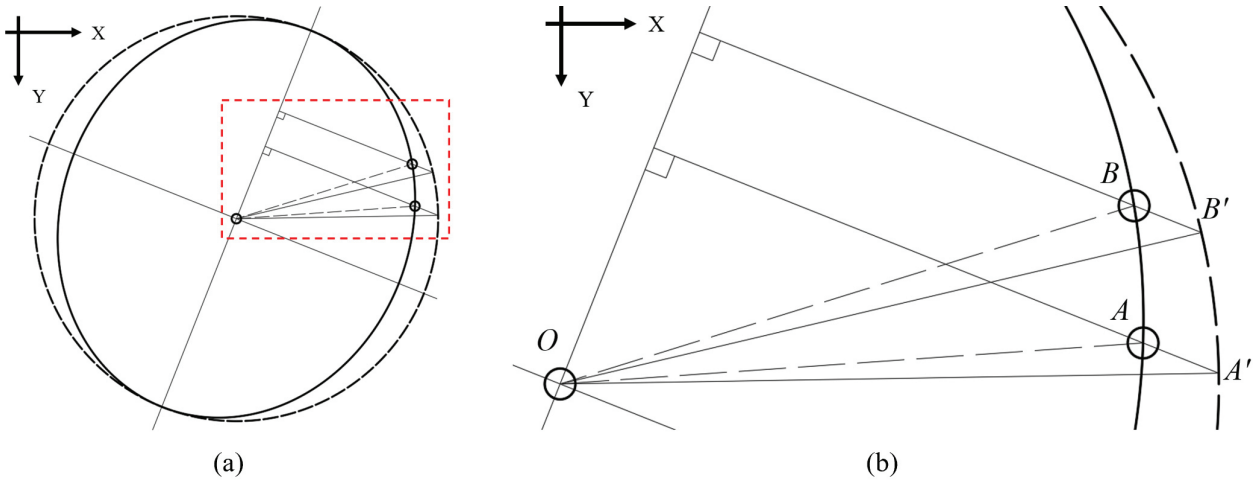


Fig. 7 (a) Schematic of the elliptical boundary and the circular boundary, and (b) a zoomed-in view of the schematic showing the transformation of the ends of the rotating blade on the elliptical boundary to the circular boundary.

angle between line *OA'* and the X-axis is

$$\beta_{A'} = \arctan\left[\frac{a}{b}\tan\left(\beta_A - \varphi\right)\right] + \varphi \tag{8}$$

Let the angle between line *OB* and the X-axis be β_B, and the angle, $\beta_{B'}$, between line OB' and the X-axis can be calculated using Eq. (8) by substituting in β_B and $\beta_{B'}$. The angle $\angle B'OA' = \beta_{A'} - \beta_{B'}$ is the angle between the two middle lines on the original circular trajectory, which is considered as the real angular movement of the rotating blade during the time period T between the two neighboring frames captured by the camera. The rotation speed of the blade in revolutions per minute (rpm) is

$$R = \frac{30(\beta_{A'} - \beta_{B'})}{\pi T} \tag{9}$$

Note that the ellipse center is not the rotation center due to the perspective projection (Figure 3), so the angles obtained by the transformation in Eq. (8) are approximations of the actual angles. These approximated angles are subsequently used to calculate the rotation speed of the blade in this work. In previous research, the rotation speed was calculated using the angle $\angle BOA$ [28], which is different from the rotation speed that is calculated when using the angle $\angle B'OA'$, and this is discussed in Sec. 3.2.

The improved demodulation method is used for OMA of the measured responses of the rotating blade [22]. For estimating the end-to-end undamped mode shapes of the rotating blade, square roots of the sum of squares of the X- and Y-mirror signals are calculated and used as position coordinates of the laser spot on a scan path during scanning of the rotating blade [22]. The measured responses are filtered by a band-pass filter that only contains one damped natural frequency of the rotating blade. The filtered responses are first multiplied by sinusoidal signals whose frequencies are the damped natural frequency, and then the responses are filtered by a low-pass filter to estimate the undamped mode shape corresponding to the damped natural frequency.

Experimental Investigation

Experimental Setup

In this work, an image-based long-range tracking CSLDV developed in Ref. [29] was used for tracking and scanning a rotating blade. The image-based long-range tracking CSLDV consists of a Polytec RSV-150 remote sensing vibrometer, a Cambridge 6240H scanner, and a Basler acA1300-200um camera (Figure 8(a)). The scanner has a set of X- and Y-mirrors that are connected to two stepper motors, and the rotation axes of the motors are perpendicular to each other. The image-based long-range tracking CSLDV is connected to an NI controller that generates input signals to the scanner and receives the measured responses and feedback signals from the scanner. One of the three aluminum blades, with a width of 0.1 m

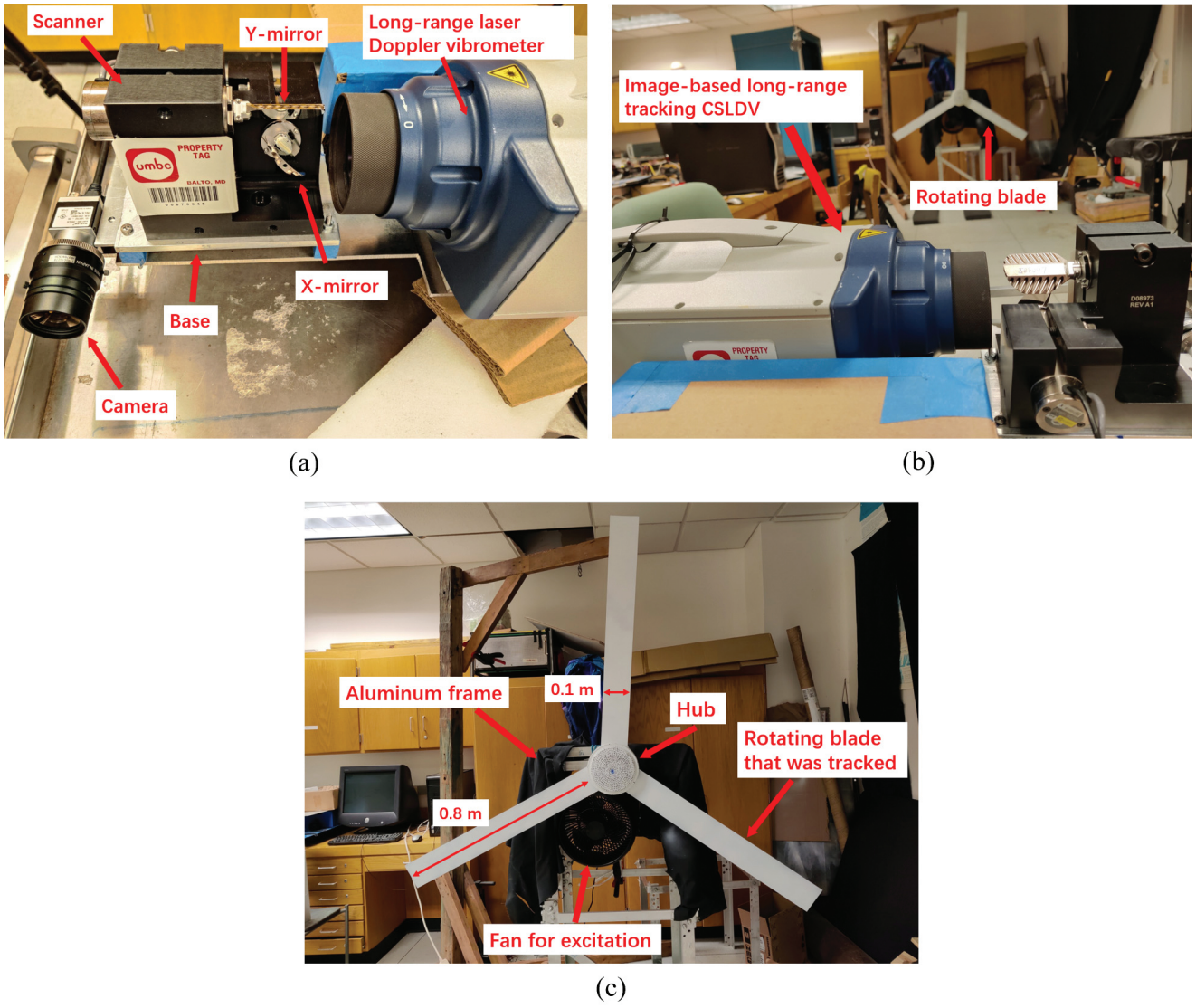


Fig. 8 (a) Photograph of the image-based long-range tracking CSLDV, (b) the experimental setup involving the tracking and scanning of the rotating blade with the image-based long-range tracking CSLDV, and (c) a photograph showing the rotating blade that was tracked.

and a length of 0.8 m, that were attached to a hub was tracked by the camera (Figure 8(b)). The hub was mounted on an aluminum frame, and the hub rotated with a constant speed when the internal motor was switched on (Figure 8(c)).

For the investigation of tracking and scanning a rotating blade with a misalignment angle between its rotation axis and the optical axis of the camera, the aluminum frame was turned at different positions so that there were different misalignment angles of 0°, 10°, 20°, and 30° between the tracking CSLDV and the rotating blade. The blade rotated with a constant speed for different misalignment angles, and the camera captured images of the rotating blade with a frame rate of 25 frames per second. Air flow from a separate fan positioned behind the rotating blades was used to excite the rotating blades. The distance between the center of the hub and the image-based long-range tracking CSLDV was 5.07 m with no misalignment angle, and the distance was 5.16 m with a misalignment angle of 30°.

Misalignment Angle Correction Method Results with Different Misalignment Angles

Captured images of the rotating blade with different misalignment angles are shown in Figure 9, where the image with no misalignment angle is shown in Figure 9(a) and images with misalignment angles of 10°, 20°, and 30° are shown in Figure 9(b), (c), and (d), respectively. Rotation trajectories of the blade with different misalignment angles are shown in

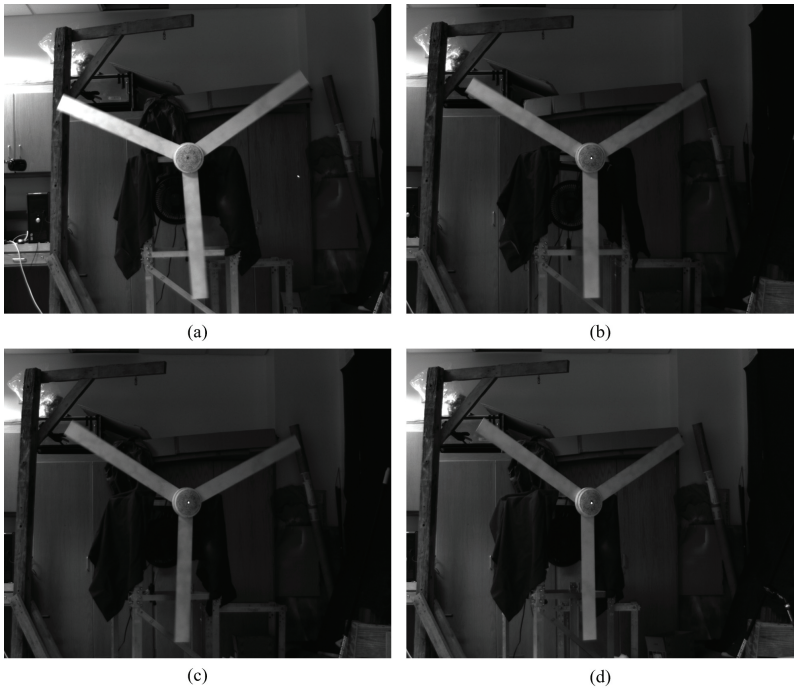


Fig. 9 Captured images of the rotating blade with (a) no misalignment angle, (b) a misalignment angle of 10°, (c) a misalignment angle of 20°, and (d) a misalignment angle of 30°.

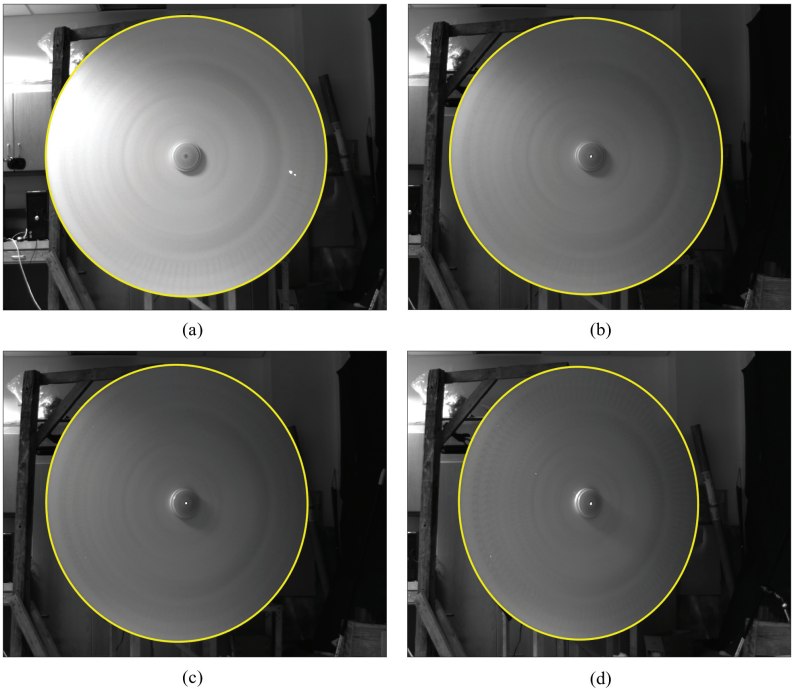


Fig. 10 Trajectories of the rotating blade with (a) no misalignment angle, (b) a misalignment angle of 10°, (c) a misalignment angle of 20°, and (d) a misalignment angle of 30°.

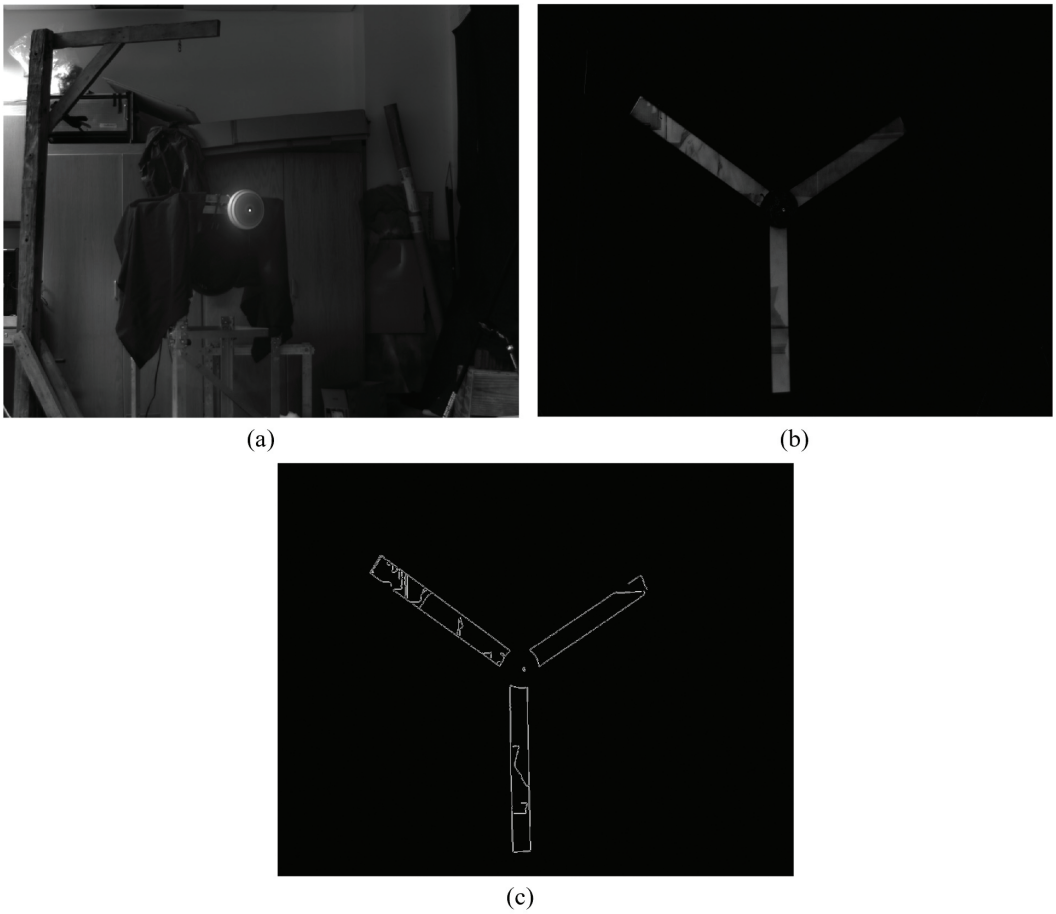
Figure 10, where each trajectory is an overlay of 500 images captured by the camera while the blade rotated with a constant speed. The boundaries of the rotation trajectories were identified as ellipses using the ellipse detection algorithm given in Ref. [30], and the parameters of the ellipses shown in Figure 10 are presented in Table 1. Note that the lengths of the major

Table 1 Estimated parameters of the elliptical boundaries of the trajectories of the rotating blade with different misalignment angles.

Misalignment Angle	a (pixels)	b (pixels)	(p, q) (pixels)	φ	θ
0°	468.61	468.44	(610.57, 509.20)	−88.54°	1.54°
10°	462.48	455.09	(594.62, 509.34)	−88.51°	10.26°
20°	462.91	436.12	(579.99, 509.57)	−88.51°	19.59°
30°	455.99	399.16	(571.26, 508.62)	−88.47°	28.91°

axis and minor axis of the ellipse in Figure 10(a) are approximately the same, meaning that the boundary is close to a circle with no misalignment angle. The angles, θ, between the rotation axis and the optical axis of the camera were calculated using Eq. (1). The rotation plane of the blade was approximately perpendicular to the ground and the aluminum frame was turned about the vertical direction. Therefore, the major axes of the resulting elliptical boundaries shown in Figure 10 are approximately vertical. Thus, the calculated angles, φ, between the major axes of the ellipses and the X-axis given in Table 1 are close to $-90°$.

The background averaging method described in Ref. [29] was used to process the captured images from the camera, and the processing of Figure 9(d) is presented here as an example. The averaged image using 500 captured images for the case with a misalignment angle of 30° is shown in Figure 11(a). To remove the background from the scene, the averaged image was subtracted from a captured image, such as the image shown in Figure 9(d). In the resulting image, the three blades remain and the background is removed, as shown in Figure 11(b). The Canny edge detection method was applied



(a) (b)

(c)

Fig. 11 (a) Averaged image using 500 captured images for the case with a misalignment angle of 30°, (b) the resulting image obtained by subtracting the averaged image from the captured image shown in Figure 9(d), and (c) the edge map obtained by applying the Canny edge detection method to (b).

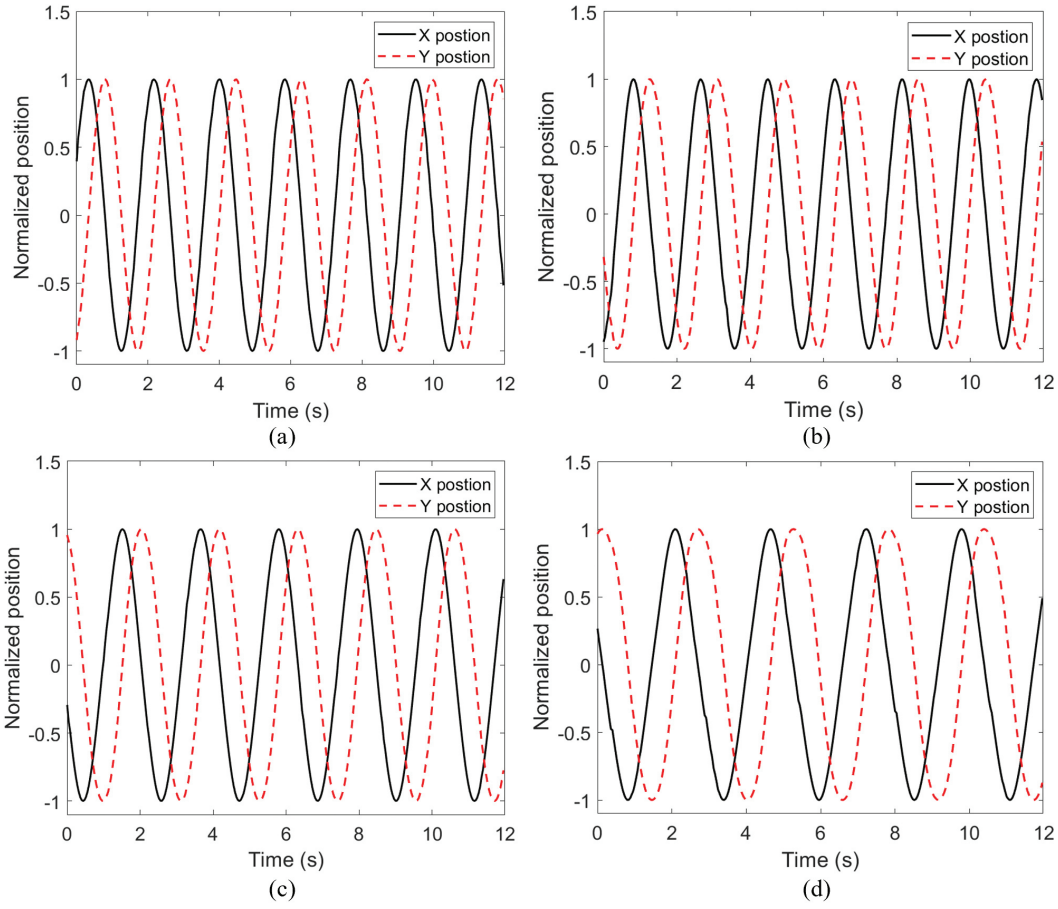


Fig. 12 Normalized position coordinates of the edge point with (a) no misalignment angle, (b) a misalignment angle of 10°, (c) a misalignment angle of 20°, and (d) a misalignment angle of 30°.

to Figure 11(b) and the resulting edge map is shown in Figure 11(c), where the edges of the three blades are depicted. For tracking one of the three rotating blades, the distance condition described in Ref. [27] was used to determine the position of an edge point on the rotating blade.

The normalized estimated position coordinates of the edge point of the rotating blade are shown in Figure 12 for each misalignment angle case, where the position coordinates of the rotation center were subtracted from the estimated position coordinates of the edge point and the resulting position coordinates were normalized by dividing by their maximum values. The estimated position coordinates were processed by using the novel method described in Sec. 2.3, where the elliptical trajectory was transformed to the circular trajectory to obtain the estimated rotation speeds of the blade at different misalignment angles. The estimated rotation speeds of the blade at different misalignment angles are shown in Figure 13. The averages of the estimated rotation speeds for each misalignment angle case shown in Figure 13 are given in Table 2, where the percent errors between the average rotation speeds with angular misalignments (10°, 20°, 30°) and the average rotation speed with no angular misalignment (0°) are calculated using the equation, $\%\ error = \frac{approximate\ value - true\ value}{true\ value} \times 100$, where the "true value" is taken as the case with no misalignment angle and the non-zero misalignment angle cases are taken as the "approximate values". For the misalignment angle of 30°, the average value of the rotation speed calculated using the legacy method is 32.78 rpm which is nearly identical to the average rotation speed of 32.79 rpm calculated using the novel method, as shown in Table 2.

OMA Results with Different Misalignment Angles

The damped natural frequencies of the rotating blade with different misalignment angles were estimated from fast Fourier transform (FFT) results of the measured responses of the rotating blade [29], and they are given in Table 3. The relative percent error between the case of no misalignment angle and the non-zero misalignment angle cases is given in Table 3.

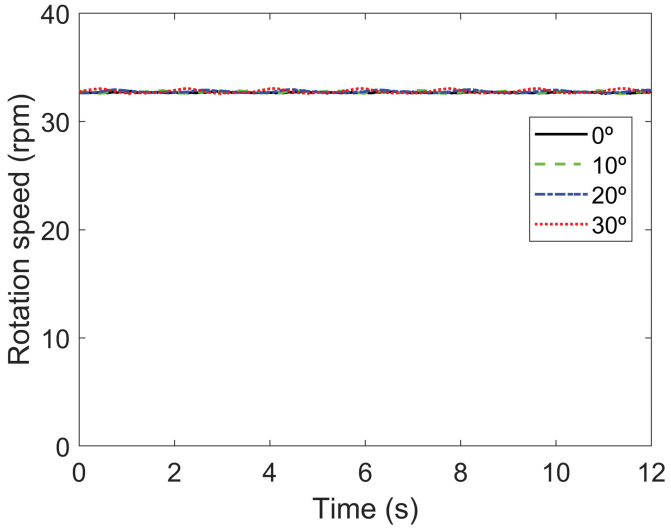


Fig. 13 Estimated rotation speeds of the blade processed using the novel method described in Sec. 2.3.

Table 2 Estimated averaged rotation speeds calculated using the novel method for different misalignment angles as well as their corresponding percentage errors.

Misalignment Angle	Average Rotation Speed (rpm)	Percent Error (%)
0°	32.68	——
10°	32.74	0.18
20°	32.77	0.28
30°	32.79	0.34

Table 3 Estimated damped natural frequencies of the rotating blade with different misalignment angles as well as their corresponding percentage errors.

Misalignment Angle	First Damped Natural Frequency (Hz)	Second Damped Natural Frequency (Hz)	Third Damped Natural Frequency (Hz)	% Error First Mode (%)	% Error % Second Mode (%)	% Error Third Mode (%)
0°	1.631	9.248	29.655	-	-	-
10°	1.655	9.269	29.887	1.47	0.23	0.78
20°	1.642	9.292	29.821	0.67	0.48	0.56
30°	1.635	9.233	29.372	0.24	-0.16	-0.96
30° without considering projection	1.646	9.295	29.434	0.92	0.51	-0.75

Note that the differences between the damped natural frequencies of the rotating blade with no misalignment angle and those with a misalignment angle are relatively small, with a maximum percent error of 1.47%. If the laser spot is swept over the surface of the rotating blade, its damped natural frequencies can be accurately estimated, independent of the tracking and scanning method used.

The first three undamped mode shapes of the rotating blade were estimated using the improved demodulation method. The estimated undamped mode shapes of the rotating blade with different misalignment angles, where the rotation speed was calculated using the novel method that considers projection in the captured images from the camera, are shown in Figure 14. These mode shapes are out-of-plane bending mode shapes of a rotating beam, and the results show that there is good agreement of the mode shapes between each angular misalignment case. The undamped mode shapes of the rotating blade with a misalignment angle of 30° were estimated using both the novel and legacy methods as shown in Figure 15. Modal assurance criterion (MAC) values between the undamped mode shapes with no misalignment angle and the undamped mode shapes with different misalignment angles are shown in Table 4. The MAC values between the estimated undamped mode shapes with no misalignment angle and those with non-zero misalignment angles calculated using the novel method that considers projection are very high (0.97 or greater), while the MAC values between the estimated undamped mode shapes with no misalignment angle and those with the misalignment angle of 30° calculated using the legacy method that does not

Table 4 MAC values between the estimated undamped mode shapes of the rotating blade with no misalignment angle and the estimated undamped mode shapes with different misalignment angles.

Angular Misalignment	First Mode	Second Mode	Third Mode
10°	0.99	0.99	0.97
20°	0.99	0.98	0.97
30°	0.99	0.99	0.97
30° without considering projection	0.99	0.94	0.89

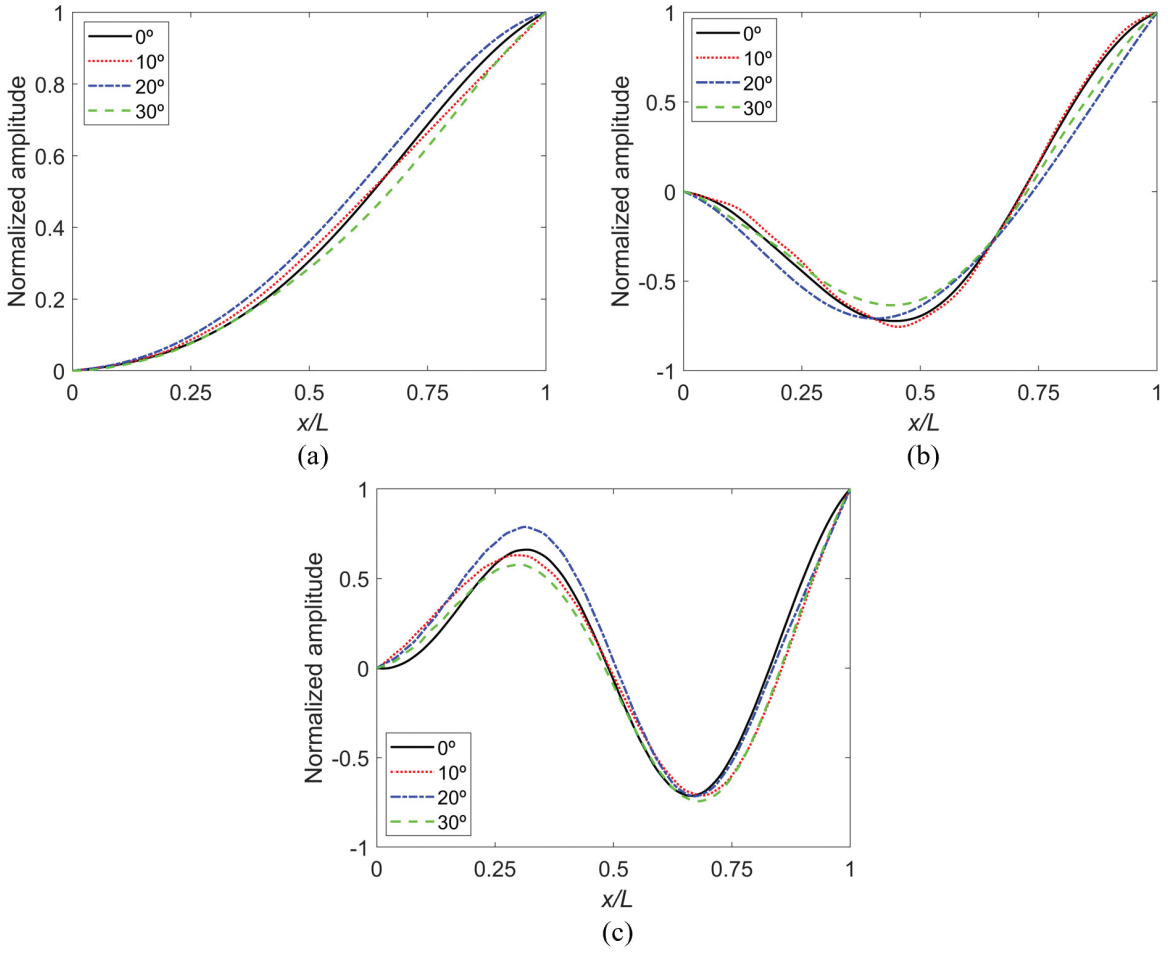


Fig. 14 Estimated (a) first, (b) second, and (c) third undamped mode shapes of the rotating blade with different misalignment angles calculated using the novel method described in Sec. 2.3.

consider projection have lower values for the second and third modes (0.94 and 0.89, respectively). Using the novel method that considers projection in the captured images can greatly reduce differences in the estimated undamped mode shapes caused by angular misalignments between the rotation axis and the optical axis of the camera.

Application of the Proposed Method on Tracking the Rotating Blade in an Outdoor Environment

In this section, the proposed method was experimentally validated through an outdoor test. As shown in Fig 16(a), the frame was moved outside the lab environment, and the motor in the hub was connected to a portable power station so that it could be switched on and rotate with a constant speed. The trajectory of the rotating blades with its boundary is shown in Figure 16(b). The semi-major axis and semi-minor axis of the boundary are 299.4 pixels and 298.9 pixels, respectively. Using Eq. (1), the misalignment angle between the rotation axis of the blades and the optical axis of the camera is 3.3°.

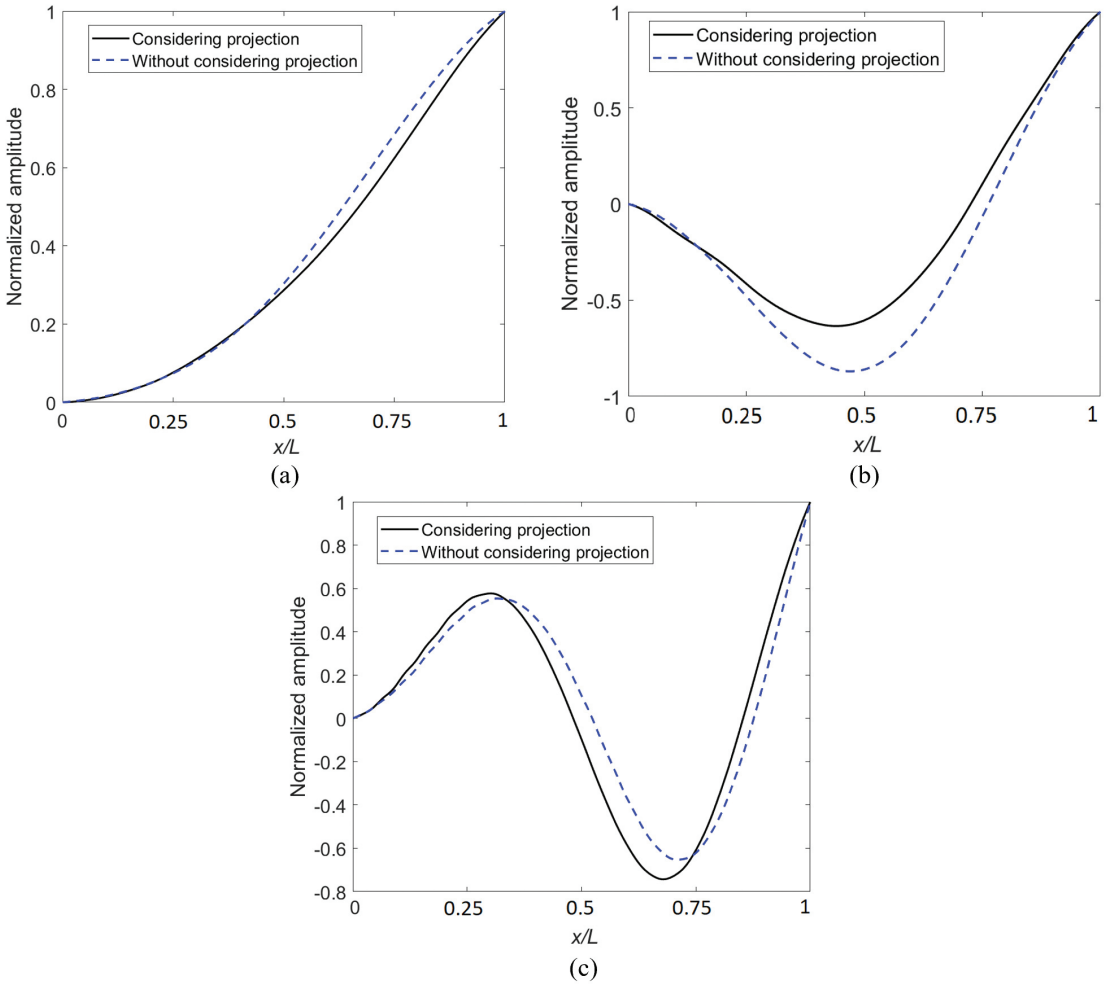


Fig. 15 Estimated (a) first, (b) second, and (c) third undamped mode shapes of the rotating blade with a misalignment angle of 30° using the novel misalignment angle correction method that considers projection and the legacy method that does not consider projection.

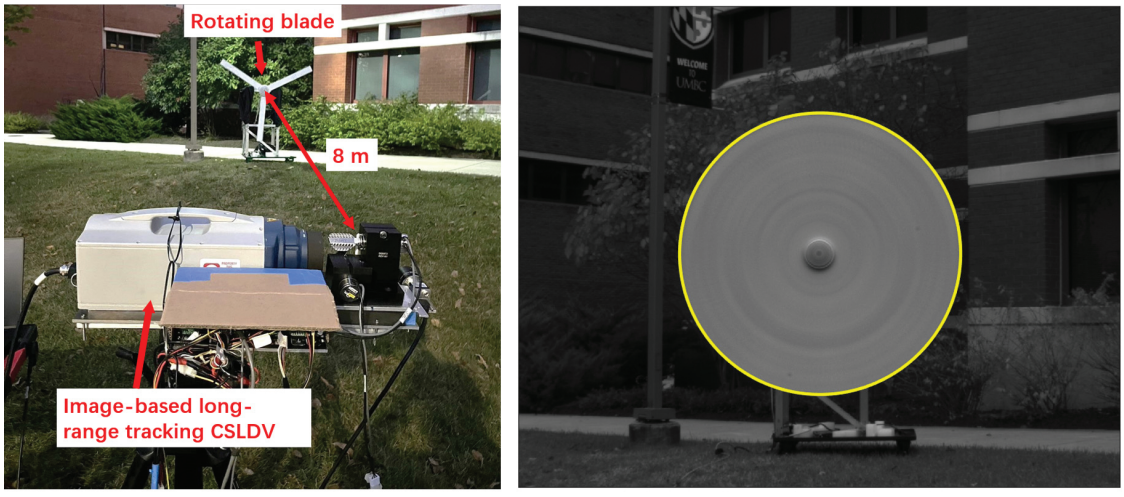


Fig. 16 (a) Experimental setup of the outdoor test for validating the proposed method of tracking and scanning a rotating blade, and (b) the estimated boundary of the trajectory of the rotating blades.

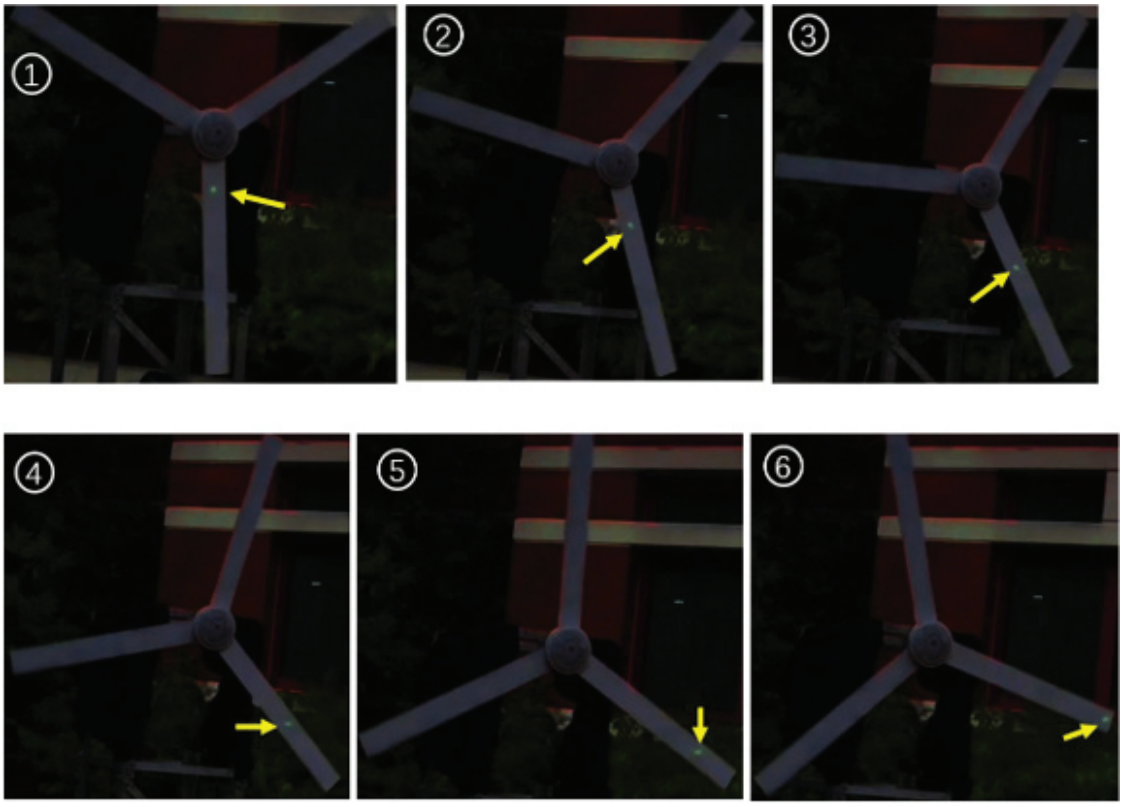


Fig. 17 Series of frames captured by the camera showing the location of the laser spot along its scan path as the blade rotated.

Straight scan paths were generated along the direction of the middle line of the rotating blade based on the pixel positions of the edge point. A series of six frames captured by the camera in the image-based long-range tracking CSLDV are shown in Figure 17, where the locations of the laser spot are indicated by arrows. Note that the frames shown in Figure 17 were processed to increase the contrast between the laser spots and the other objects in the frames. In the series of frames, the laser spot is moved along the middle line of the rotating blade, which is its generated scan path used for capturing vibration data of the rotating structure.

Conclusion

A novel misalignment angle correction method was developed for processing images of a rotating structure with a large misalignment angle between its rotation axis and the optical axis of the camera that captured its images. The rotation trajectory of the structure is considered as an ellipse due to the projection in the captured images from the camera. Parameters of the ellipse, including the position of its center, the lengths of its semi-major axis and semi-minor axis, and the angular position of its major axis, are used to determine the misalignment angle and the endpoints of the scan paths on the rotating structure. The difference between the center of the ellipse and the rotation center of the structure is considered in the misalignment angle correction method to determine accurate position coordinates of the rotating structure in the captured images. The estimated position coordinates of the end of the rotating structure in the elliptical trajectory are transformed into position coordinates in the original circular rotation trajectory to obtain the estimated rotation speeds of the blade at different misalignment angles.

An aluminum blade attached to a hub was used as a rotating structure for the experimental investigation of the misalignment angle correction method. The rotating blade with different misalignment angles of 0°, 10°, 20°, and 30° was tracked and scanned by the tracking CSLDV and its modal parameters, including the damped natural frequencies and undamped mode shapes, were estimated by the improved demodulation method. The estimated damped natural frequencies of the rotating blade with different misalignment angles matched well, with a maximum percent error of 1.47% when comparing the case of no misalignment angle to the non-zero misalignment angle cases, independent of the method used. The estimated

undamped mode shapes of the rotating blade with a large misalignment angle calculated using the novel misalignment angle correction method showed that there is good agreement with the undamped mode shapes of the rotating blade with no misalignment angle.

In addition to the presented results, this study highlights several key aspects of the proposed misalignment angle correction method. One of the primary advantages is that the method enables accurate scan path generation for rotating structures with significant misalignment angles, without requiring additional physical markers or encoders. This makes it well-suited for non-invasive applications in real-world scenarios. However, limitations exist in terms of hardware constraints, such as the field of view of the camera and the working distance of the LDV. While the method itself is theoretically applicable to misalignment angles of greater than 30°, practical implementation beyond this range would require more advanced imaging hardware and tracking algorithms. These findings suggest that the method has strong potential for application in full-scale rotating systems, provided that environmental and optical conditions are properly managed.

Acknowledgments This research was supported by the National Science Foundation through Grant No. CMMI-1763024.

References

1. S. Rothberg, M. Allen, and P. Castellini, An international review of laser Doppler vibrometry: making light work of vibration measurement, *Opt. Lasers Eng.* 99(1) (2017), pp. 11–22.
2. B. Stoffregen, and A. Felske, Scanning laser Doppler vibration analysis system. *SAE Transactions*, 94(2) (1985), pp 934-940.
3. P. Castellini, M. Martarelli, and E. Tomasini, Laser Doppler vibrometry: development of advanced solutions answering to technology's needs. *Mech. Syst. Signal Process*. 20(6) (2006) pp. 1265–1285.
4. K. Yuan, and W. Zhu, Modeling of welded joints in a pyramidal truss sandwich panel using beam and shell finite elements. *J. Vib. Acoust*. 143 (4) (2021), p. 041002.
5. C. Vuye, S. Vanlanduit, F. Presezniak, G. Steenackers, P. Guillaume, Optical measurement of the dynamic strain field of a fan blade using a 3D scanning vibrometer. *Opt. Lasers Eng*. 49 (7) (2011), pp. 988-997.
6. P. Sriram, S. Hanagud, J. Craig, and N.M. Komerath, Scanning laser Doppler technique for velocity profile sensing on a moving surface, *Appl. Opt.* 29(16) (1990), pp. 2409-2417.
7. P. Sriram, S. Hanagud, and J. Craig, Mode shape measurement using a scanning laser Doppler vibrometer, *Int. J. Anal. Exp. Modal anal*. 7(3) (1992), pp. 169–178.
8. M.S. Allen, and M.W. Sracic, A new method for processing impact excited continuous-scan laser Doppler vibrometer measurements, *Mech. Syst. Signal Process.* 24(3) (2010), pp. 721–735.
9. D.M. Chen, Y.F. Xu, and W.D. Zhu, Damage identification of beams using a continuously scanning laser Doppler vibrometer system, *J. Vib. Acoust.* 138(5) (2016), p. 05011.
10. D.M. Chen, and W.D. Zhu, Investigation of three-dimensional vibration measurement by three scanning laser Doppler vibrometers in a continuously and synchronously scanning mode. *J. Sound Vib.*, 498 (2021), p. 115950.
11. K. Yuan, and W.D. Zhu, Estimation of modal parameters of a beam under random excitation using a novel 3D continuously scanning laser Doppler vibrometer system and an extended demodulation method, *Mech. Syst. Signal Process.* 155 (2021), p. 107606.
12. K. Yuan, and W.D. Zhu, In-plane operating deflection shape measurement of an aluminum plate using a three-dimensional continuously scanning laser Doppler vibrometer system, *Exp. Mech.*, 62 (2022), pp. 667–676.
13. K. Yuan, and W.D. Zhu, A novel general-purpose three-dimensional continuously scanning laser Doppler vibrometer system for full-field vibration measurement of a structure with a curved surface, *J. Sound Vib.*, 540 (2022), p. 117274.
14. K. Yuan, and W.D. Zhu, Identification of modal parameters of a model turbine blade with a curved surface under random excitation with a three-dimensional continuously scanning laser Doppler vibrometer system, *Measurement*, 2023, p. 112759.
15. A. Stanbridge, and D. Ewins, Modal testing using a scanning laser Doppler vibrometer, *Mech. Syst. Signal Process.* 13(2) (1999), pp. 255–270.
16. D. Di Maio, and D. Ewins, Continuous scan, a method for performing modal testing using meaningful measurement parameters; Part I, *Mech. Syst. Signal Process.* 25(8) (2011), pp. 3027–3042.
17. D.M. Chen, Y.F. Xu, and W.D. Zhu, Experimental investigation of notch-type damage identification with a curvature-based method by using a continuously scanning laser Doppler vibrometer system, *J. Nondestruct. Eval.* 36(2) (2017), p. 38.
18. D.M. Chen, Y.F. Xu, and W.D. Zhu, Identification of damage in plates using full-field measurement with a continuously scanning laser Doppler vibrometer system, *J. Sound Vib.* 422 (2018), pp. 542–567.
19. D.M. Chen, Y.F. Xu, and W.D. Zhu, Non-model-based identification of delamination in laminated composite plates using a continuously scanning laser Doppler vibrometer system. *J. Vib. Acoust.* 140(4) (2018), p. 041001.
20. D.M. Chen, Y.F. Xu, and W.D. Zhu, A comprehensive study on detection of hidden delamination damage in a composite plate using curvatures of operating deflection shapes, *J. Nondestruct. Eval.* 38 (2019), p. 54.
21. Z. Huang, and C. Zang, Damage Detection Using Modal Rotational Mode Shapes Obtained with a Uniform Rate CSLDV Measurement, *Appl. Sci.* 9(23) (2019), p. 4982.
22. L.F. Lyu, and W.D. Zhu, Operational modal analysis of a rotating structure subject to random excitation using a tracking continuously scanning laser Doppler vibrometer via an improved demodulation method, *J. Vib. Acoust.*, 144 (1) (2022), p. 011006.
23. L.F. Lyu, and W.D. Zhu, Full-field mode shape estimation of a rotating structure subject to random excitation using a tracking continuously scanning laser Doppler vibrometer via a two-dimensional scan scheme, *Mech. Syst. Signal Process.* 169 (2022), p. 108532.

24. S., Yang, and M.S., Allen, Lifting approach to simplify output-only continuous-scan laser vibrometry, *Mech. Syst. Signal Process.* 45(2) (2014), pp. 267–282.
25. Y.F. Xu, D.M. Chen, and W.D. Zhu, Operational modal analysis using lifted continuously scanning laser Doppler vibrometer measurements and its application to baseline-free structural damage identification, *J. Vib. Control.* 25(7) (2019), pp. 1341-1364.
26. L.F. Lyu, and W.D. Zhu, Operational modal analysis of a rotating structure under ambient excitation using a tracking continuously scanning laser Doppler vibrometer system, *Mech. Syst. Signal Process.*, 152 (2021), p. 107367.
27. L.F. Lyu, G.D. Higgins, and W.D. Zhu, Operational modal analysis of a rotating structure using image-based tracking continuously scanning laser Doppler vibrometry via a novel edge detection method, *J. Sound Vib.*, 525 (2022), p. 116797.
28. L.F. Lyu, G.D. Higgins, and W.D. Zhu, A novel robust edge detection method for full-field modal parameter estimation using an image-based tracking continuously scanning laser Doppler vibrometer system. *Mech. Syst. Signal Process*., 195 (2023), p. 110242.
29. L. Lyu, and W. Zhu, A novel image-based long-range continuously scanning laser doppler vibrometer for operational modal analysis of a rotating structure. *Res. Nondestruct. Eval.,* (2024), pp. 1-16.
30. C. Lu, S. Xia, M. Shao, and Y. Fu, Arc-support line segments revisited: An efficient high-quality ellipse detection. *IEEE Trans. Imag. Process.,* 29 (2019), pp. 768-781.

Chapter 9
Optical Motion Magnification for Vibration-based Condition Monitoring of Hydroelectric Turbines

Hussein Ali Ahmad, Marco Civera, Cecilia Surace, Christopher Niezrecki, and Alessandro Sabato

Abstract Hydropower plants are often viewed as environmentally sustainable and economically efficient sources of energy, characterized by low operating costs and the ability to support grid stability through reservoir storage. However, the long-term reliability and performance of these facilities heavily depend on effective monitoring of critical components, particularly hydroelectric turbines. Traditional condition monitoring methods often rely on contact or line-of-sight sensors, which can be intrusive, costly, or difficult to install in certain environments. This study investigates the use of Optical Motion Magnification (OMM) as a novel approach for monitoring the dynamic behavior of turbine components. OMM works by amplifying subtle motions in high-resolution video footage, allowing visualization of fine-scale vibrations that are otherwise imperceptible to the human eye. This method enables the detection of early signs of mechanical degradation without physically interacting with the equipment. A series of laboratory-based experiments were conducted to evaluate the effectiveness of OMM in comparison to conventional non-contact sensors, including proximity probes and laser displacement systems. Key performance metrics such as sensitivity, accuracy, and fault detection capability were analyzed. The results indicate that OMM successfully captures the essential features of vibrational behavior, making it a promising alternative for condition monitoring in hydropower settings. Due to its adaptability, non-invasiveness, and potential for lower implementation costs, OMM offers a compelling option for augmenting or replacing traditional sensing technologies.

Keywords Optical Motion Magnification · Condition Monitoring · Computer Vision · Hydroelectric Turbines · Non-contact Sensors

Introduction

Condition monitoring (CM) plays a critical role in ensuring the reliability, efficiency, and longevity of rotating machinery across various industries, including power generation [1], manufacturing [2], and transportation [3]. In the context of hydropower facilities, turbine systems are particularly vital, as they operate under high loads and harsh environmental conditions for extended periods [4]. Failures in hydropower turbine components can lead to costly downtime, reduced efficiency, and even catastrophic damage, making predictive maintenance and early fault detection essential [5]. To address this, CM systems are commonly employed to detect abnormalities in machine behavior before they escalate into major failures.

Traditional CM approaches often rely on contact-based sensors such as accelerometers, proximity probes, and laser displacement sensors to extract global changes in the dynamic properties of the inspected asset [6, 7]. While effective in many applications, these technologies come with notable limitations. Their installation can be invasive, requiring physical access to critical components that may be difficult to reach or housed in sealed enclosures. Additionally, these sensors can be expensive, prone to wear, and require regular maintenance. In hydropower plants, where space constraints, fluid exposure,

Hussein Ali Ahmad · Marco Civera(✉) · Cecilia Surace
Department of Structural, Geotechnical and Building Engineering, Politecnico di Torino, 10129 Turin, Italy
e-mail: s313580@studenti.polito.it; marco.civera@polito.it; cecilia.surace@polito.it

Christopher Niezrecki · Alessandro Sabato
Department of Mechanical and Industrial Engineering, University of Massachusetts Lowell, Lowell, MA 01854, United States
e-mail: Christopher_Niezrecki@uml.edu; alessandro_sabato@uml.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

and operational complexity pose significant challenges, these limitations can compromise the feasibility and effectiveness of conventional sensing solutions.

To overcome these barriers, this study explores the use of Optical Motion Magnification (OMM) as a non-contact, vision-based alternative for monitoring the dynamic behavior of hydropower turbine components. OMM amplifies small motions in high-resolution video recordings, making otherwise imperceptible vibrations visible and measurable [8, 9]. This technique offers the potential to detect early-stage mechanical faults in a non-invasive and cost-effective manner, without disrupting equipment operation.

The objective of this study is to evaluate the feasibility and performance of OMM for condition monitoring in hydropower applications. Through a series of controlled experiments, the dynamic response of turbine components is analyzed using OMM and compared against conventional sensor outputs. The findings aim to determine whether OMM can serve as a viable supplement, or even an alternative, to traditional monitoring methods, particularly in environments where sensor placement is burdensome or intrusive. This paper is organized as follows. Section 2 (Related studies) summarizes the most used techniques for CM of rotating machinery and some of the applications of OMM. Section 3 (Laboratory experiments) describes the tests performed to validate the accuracy of OMM in measuring the vibration of rotating machinery in a controlled environment. Section 4 (Field tests) provides an overview of the results obtained when OMM is used to measure the displacement of an operational hydropower turbine. Finally, Section 5 (Conclusions) summarizes the main findings of this study and provides insight into future directions.

Related Studies and Research Context

Overview of condition monitoring techniques for rotating machinery

Rotating machinery, such as motors, pumps, compressors, turbines, and gearboxes, is integral to many industrial applications, particularly in hydropower plants, where unexpected breakdowns can result in significant production losses, safety hazards, and increased maintenance costs. Over the decades, various techniques have been developed to monitor the health of such equipment. Vibration analysis remains one of the most established and widely used CM techniques for rotating machinery [10, 11]. It relies on sensors such as accelerometers or velocity transducers mounted on key components (e.g., bearings, casings) to detect abnormal vibration signatures. The underlying principle is that mechanical defects, such as imbalance, misalignment, bearing faults, or gear damage, induce characteristic changes in vibration frequency and amplitude. Signal processing techniques are used to analyze these vibrations and identify patterns or changes in those patterns that are representative of the onset of damage [12]. Vibration analysis offers high sensitivity and diagnostic capability, but it requires careful sensor placement, experience in interpreting spectral patterns, and often necessitates shutdown for accurate assessments in some instances.

Ultrasound CM is another popular approach that uses high-frequency acoustic waves to detect anomalies, particularly crack formation in bearings [13]. Ultrasonic detectors can sense friction, turbulence, and impact-generated waves in the 20 kHz–100 kHz range. These signals often precede the onset of detectable vibration or thermal anomalies, making ultrasound highly effective for early detection of bearing lubrication issues, arcing in electrical components, and valve leakages [14]. The technique is particularly effective in high-noise environments and offers both airborne and contact modes. However, interpreting ultrasonic signals can be less intuitive and typically requires skilled personnel. Similarly to ultrasound, to ultrasound, acoustic emission (AE) monitoring involves capturing high-frequency elastic waves generated by the rapid release of energy from localized sources within materials. In rotating machinery, AE can detect crack initiation, corrosion, and material deformation [15]. AE sensors are extremely sensitive and capable of identifying microscopic defects that are often undetectable by other methods. AE techniques are particularly useful for real-time monitoring of critical components under load. However, they are susceptible to ambient noise and require signal filtering and pattern recognition techniques for reliable fault classification.

To conclude, Infrared thermography (IRT) detects temperature anomalies on the surface of machinery by capturing emitted infrared radiation. Faults such as misalignment, imbalance, and frictional heating often manifest as localized hotspots, which can be visually identified in thermal images [16]. IRT is non-contact, fast, and suitable for live monitoring without disrupting operations. It is especially effective for electrical systems, lubrication faults, and misaligned couplings [17], while its applications in the hydropower domain are quite limited. Additionally, its effectiveness is limited in identifying faults that do not result in significant thermal changes, and environmental factors (e.g., airflow, emissivity variations) can impact accuracy. For a detailed overview of CM techniques for hydropower turbines, interested readers can refer to [18, 19]. These traditional techniques, while effective, often require complex setups, high expertise, and in some cases, contact-based sensing, which may not always be feasible. This has led to increased interest in non-contact, image-based techniques such as optical motion magnification [20, 21].

Principles and Applications of Optical Motion Magnification

Optical motion magnification (OMM) is a computer vision technique that visually enhances subtle motions in video recordings, making them perceptible to the human eye or measurable through image processing. Unlike traditional methods, OMM enables full-field visualization of structural or mechanical vibrations without physical sensors, allowing safe, remote diagnostics and intuitive fault detection [22].

OMM techniques can be broadly categorized into two methodological frameworks: Eulerian and Lagrangian [23]. The Eulerian approach analyzes intensity changes at fixed spatial locations over time. Each pixel in a video is considered a temporal signal, and motion is inferred from periodic changes in intensity caused by subtle displacements. A temporal band-pass filter is applied to isolate specific frequency bands corresponding to mechanical vibrations. The resulting filtered signals are amplified and recombined with the original video, producing a visually enhanced version of motion. Eulerian techniques are computationally efficient and particularly effective for small, linear displacements. In contrast, the Lagrangian approach tracks the displacement of features or pixels over time. Optical flow or block matching algorithms are used to estimate the trajectory of points, and the motion vectors are amplified before rendering the output. This method is well-suited for larger or non-linear deformations but is generally more computationally intensive and sensitive to tracking errors or noise.

The general OMM workflow begins with the acquisition of a high-frame-rate video of the target machinery. The video is decomposed either spatially (e.g., using Laplacian pyramids or wavelets) and temporally to extract motion information. The filtered signals are then amplified by a magnification factor α, which depends on noise sensitivity and the desired visibility [9]. The amplified signals are added back to the original video frames to create an output that visually exaggerates the motion while preserving the spatial structure. Key parameters such as magnification factor, temporal bandwidth, and filter type must be carefully chosen to prevent artifacts, maintain fidelity, and avoid ghosting [24]. Examples of applications of OMM include operational modal analysis [25–29], monitoring of structures with cracks [30], vibrating panels [31], as well as wind turbine foundations [32, 33]. An evaluation of the performance of this method for CM of hydropower turbines is missing and motivates this study.

Laboratory Experiments

To assess the precision and feasibility of video-based vibration measurement for CM of rotating machinery, a series of laboratory experiments were conducted. The objective was to evaluate the OMM accuracy in capturing the vertical displacements of an axle test rig with varying off-axis mass configurations, and to compare its performance with that of conventional sensing methods. In the experiments, the test rig, shown in Figure 1a, was deliberately unbalanced by attaching masses to the rotating shaft. The resulting dynamic response was simultaneously measured using a high-speed camera and a laser displacement sensor. Masses ranging from 0 to 8×10^{-3} kg were incrementally added in steps of 2×10^{-3} kg, and the rig was operated at a constant rotational speed of 2,400 rpm. Video recordings were captured using a 5-megapixel IRIS MS camera equipped with a 25mm lens, positioned approximately 0.5 meters from the setup, and acquiring images at 1,400 frames per second (fps). The collected videos were processed using the software Motion Amplification, developed by RDI Technologies, to extract both the vertical displacement of the shaft and the system's frequency domain response under each test condition. Concurrently, a Keyence LK-G402 laser displacement sensor was employed to acquire ground-truth vertical displacement data at a sampling rate of 1,000 Hz. An overview of the experimental setup is provided in Figure 1b. To ensure statistical robustness, each mass-loading scenario was tested three times. An example of the time histories recorded by the camera-based system and the laser distance meter is presented in Figure 2a, while Figure 2b displays the corresponding frequency domain response obtained using a Fast Fourier Transform (FFT). As shown, the two signals exhibit strong correlation, with peak-to-peak amplitudes of 26.81 ± 1 and 25.43 ± 2.02 for the camera and laser measurements, respectively. In the frequency domain, both signals reveal a distinct and consistent peak at 40 Hz, corresponding to the imposed rotational speed of 2,400 rpm, further confirming the agreement between the two sensing modalities.

Since each test was repeated three times, it was possible to compute the mean and standard deviation of the measured amplitudes from both systems, enabling a quantitative comparison of their performance. These results are summarized in Figure 3. As observed, the measurements obtained from the camera-based system using OMM closely match those from the laser sensor, with all data points falling within one standard deviation of difference. This consistency validates the accuracy and reliability of the camera-based method in replicating the reference measurements provided by the laser displacement sensor. Since the system developed in this study is ultimately intended for inspecting operational hydropower turbines, where close camera placement may not be feasible, a second laboratory experiment was conducted to evaluate the performance of OMM as a function of working distance, field of view (FOV), and the camera's noise floor. In this experiment, the noise floor

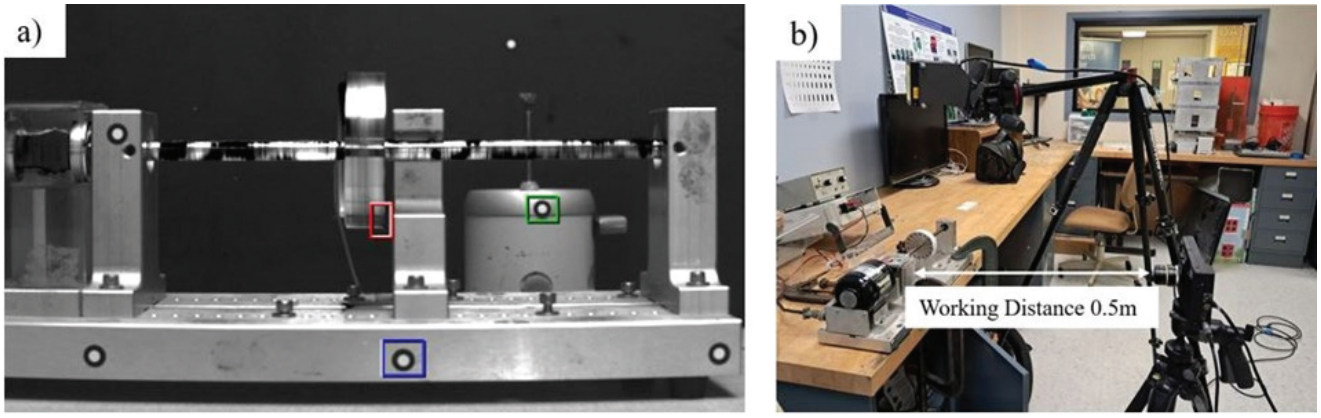


Fig. 1 Overview of the laboratory experiments used to validate the accuracy of OMM compared to traditional laser-based measurements: a) axle test rig as viewed by the IRIS MS camera and b) experimental setup.

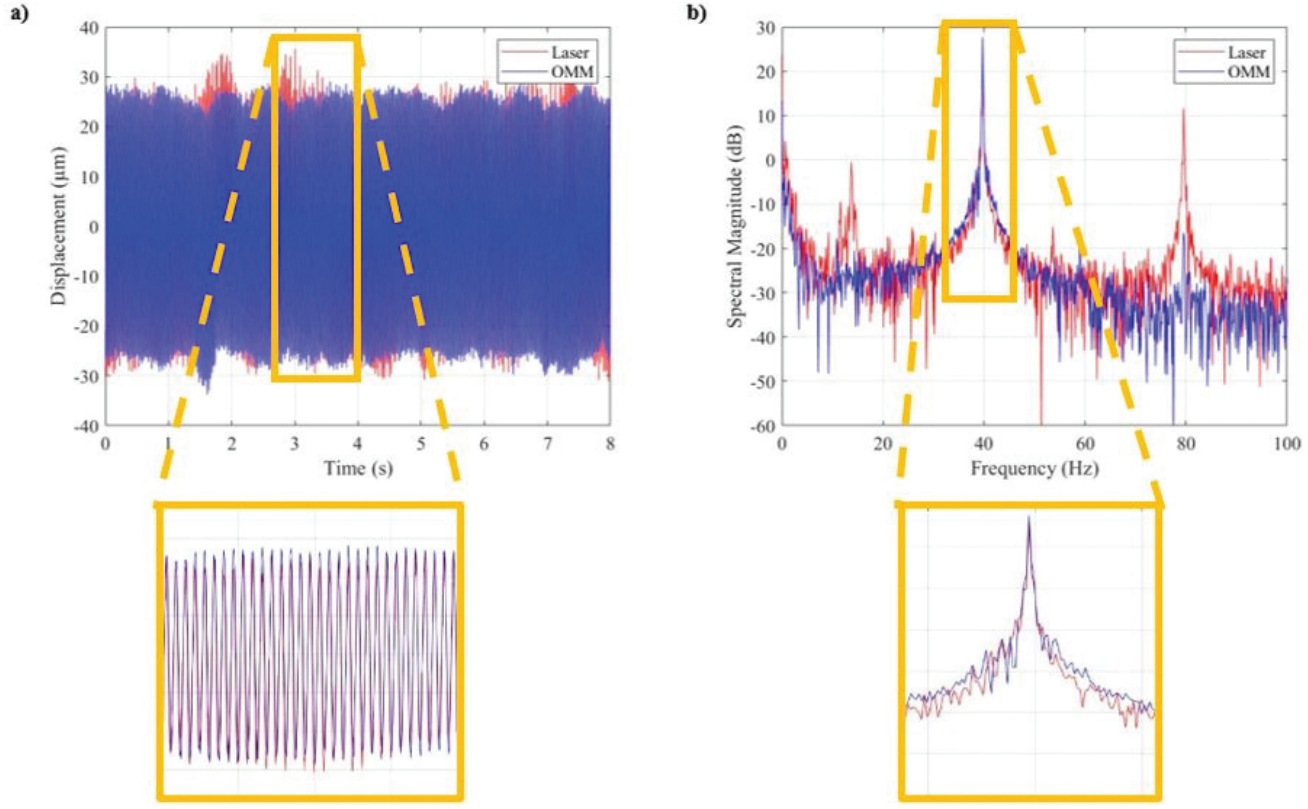


Fig. 2 Example of displacement signals in the Y-direction collected with the laser displacement sensor and the camera-based system for the case in which a mass of $6 \cdot 10^{-3}$ kg was added to the axle test rig: a) time domain results and b) corresponding frequency spectra.

of the camera was quantified by imaging a static, non-vibrating scene at working distances ranging from 1 to 7 meters, in 1-meter increments. The resulting measures, which provide insight into how signal quality degrades with increasing distance, are presented in Figure 4.

In this context, the noise floor is defined as the root mean square (RMS) of the signal measured along the x- and y-axes. Under ideal conditions, given that the camera images a stationary target, the measured displacement should be zero. However, due to inherent sensor noise, the camera produces displacement readings typically ranging between 5 and 40 μm. As expected, the noise floor of the camera-based system increases with working distance, which also corresponds to an increase in the field of view (FOV) and a reduction in spatial resolution.

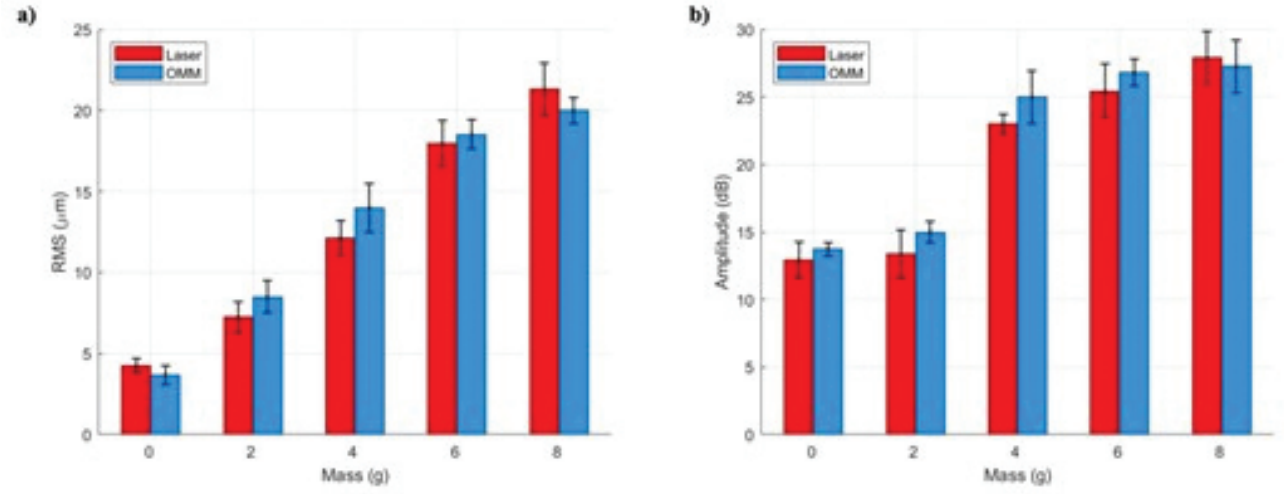


Fig. 3 Comparison of results obtained in the first laboratory experiment for all the case studies considered: a) time domain comparison and b) frequency domain comparison.

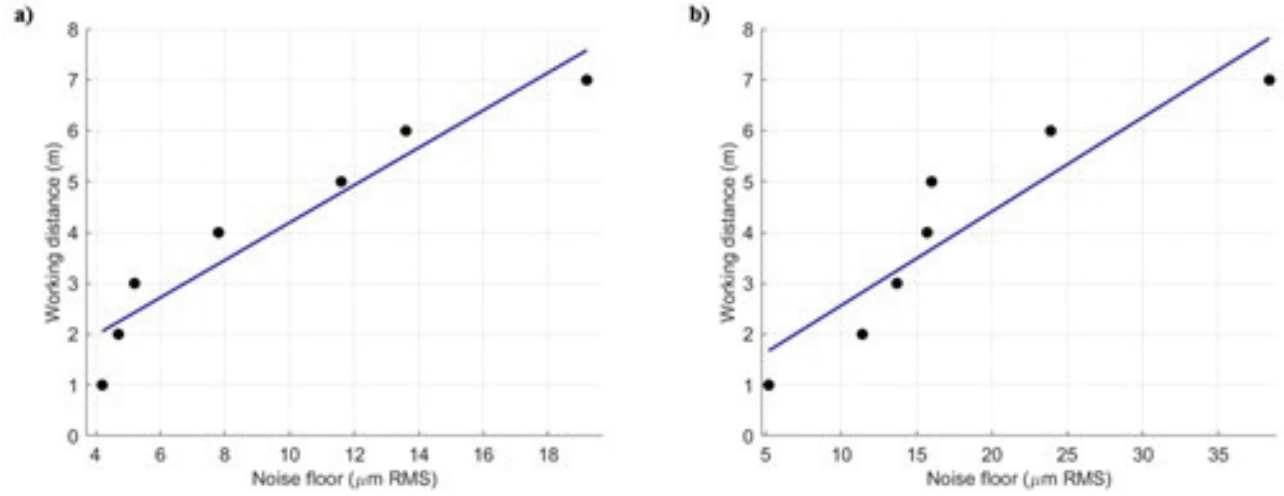


Fig. 4 IRIS MX's noise floor as a function of the working distance expressed in terms of the signal's RMS in the a) X-direction and b) Y-direction.

To assess the viability of the OMM system for measuring displacement at working distances beyond 1 meter, a third laboratory experiment was conducted. In this test, a target object was mounted on a shaker and excited with a 3 Hz sinusoidal input. The vertical displacement of the object was simultaneously measured using both the camera-based system and a laser displacement sensor. Measurements were collected at working distances ranging from 1 to 7 meters, in 1-meter increments. To ensure statistical robustness, each test condition was repeated three times. Representative results obtained at 1 m and 7 m working distances are illustrated in Figure 5.

As shown in Figure 5a, the time series at a working distance of 1 meter reveals strong agreement between the camera-based system and the laser displacement sensor, with nearly indistinguishable waveforms. This close match indicates that the camera-based system is capable of accurately replicating the reference measurements at short range. Figure 5b provides a quantitative comparison using the Time Response Assurance Criterion (TRAC), confirming a correlation above 99% between the two signals. At a working distance of 7 meters, the results in Figure 5c show a somewhat reduced level of agreement; nevertheless, the camera-based system still successfully captures the displacement of the target. In this case, Figure 5d shows a TRAC value of approximately 95%, with the scatter plot exhibiting greater dispersion around the unity line, indicating a modest decline in measurement accuracy. Despite this reduction, the results confirm that OMM remains a viable and reliable

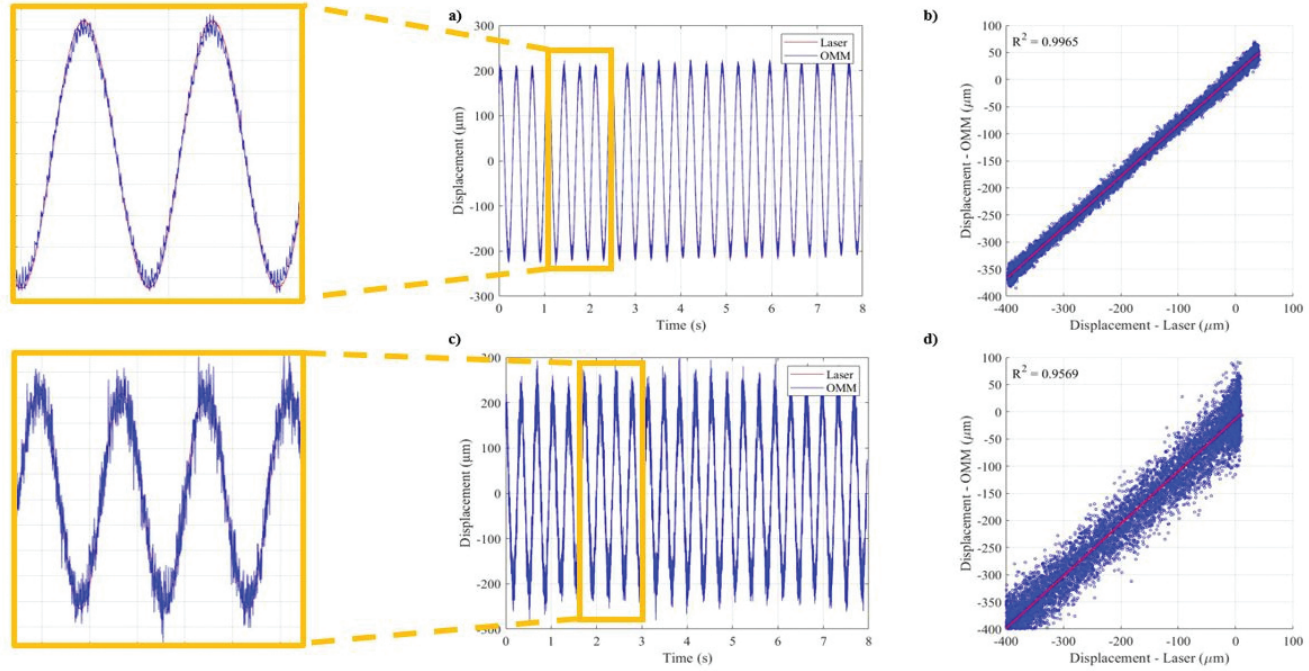


Fig. 5 Comparison of signals measured with the camera-based system at different working distances against laser distance sensors: a) time series at 1 m working distance, b) corresponding TRAC between the two datasets, c) time series at 7 m working distance, and d) corresponding TRAC between the two datasets.

technique for displacement measurement even at extended working distances, with only a minor degradation in performance compared to the reference sensor.

Field Tests

The field tests described in this paper were conducted at the Boott Hydropower Plant in Lowell, MA, on an 8.65 MW horizontal Francis turbine manufactured by Fuji Electric (see Figure 6a). The turbine operates with a head of ∼12 m and a rotational speed of 120 rpm (i.e., 2 Hz), creating a dynamic yet controlled environment for experimental validation. The goal of these tests was to validate OMM's accuracy and reliability in measuring the displacement of turbine components under real operating conditions and to compare its performance with that of conventional sensors, such as laser displacement sensors and proximity probes.

To collect the necessary data, three different sensing modalities were deployed: the 5-megapixel IRIS MX camera operating at 125 frames per second, the Keyence LK-G402 laser displacement sensor sampling at 1,000 Hz, and pre-installed proximity probes on the turbine shaft. As shown in Figure 6b and Figure 6c, the camera was positioned at two distinct working distances: 2.70 meters from Shaft 1 and 2.27 meters from Shaft 2. In both tests, the camera was used to measure the vertical displacement (i.e., Y-direction) of the two shafts as they vibrated during turbine operation. The proximity probes, installed on the turbine by the plant's management company, are used to record displacements along both the Y- and Z-directions for Shaft 1 and the Y-direction for Shaft 2. Similarly, the laser sensor mounted near Shaft 1 was used to measure the displacement of this component in the Z-direction (see Figure 6d). Each configuration was tested twice to ensure the reproducibility of results and enable statistical analysis.

The results of the tests are shown in Figure 7. In Figure 7a, the data obtained for Shaft 1 show that the average peak-to-peak displacement in the Y-direction measured using OMM was 162 ±8.5 μm. This value was corroborated by the reading of the proximity probe measurements, which showed close agreement, confirming the camera's capability to resolve motion at these distances with acceptable precision. These results were visualized using a series of images comparing different magnification levels (α = 0, 50, and 100), emphasizing how motion becomes more visually interpretable with increased magnification. As shown in Figure 7b, the OMM analysis produced an average peak-to-peak Y-displacement of 129±3.5 μm. This value was likewise validated against proximity probe data, with a minimal difference between the two methods.

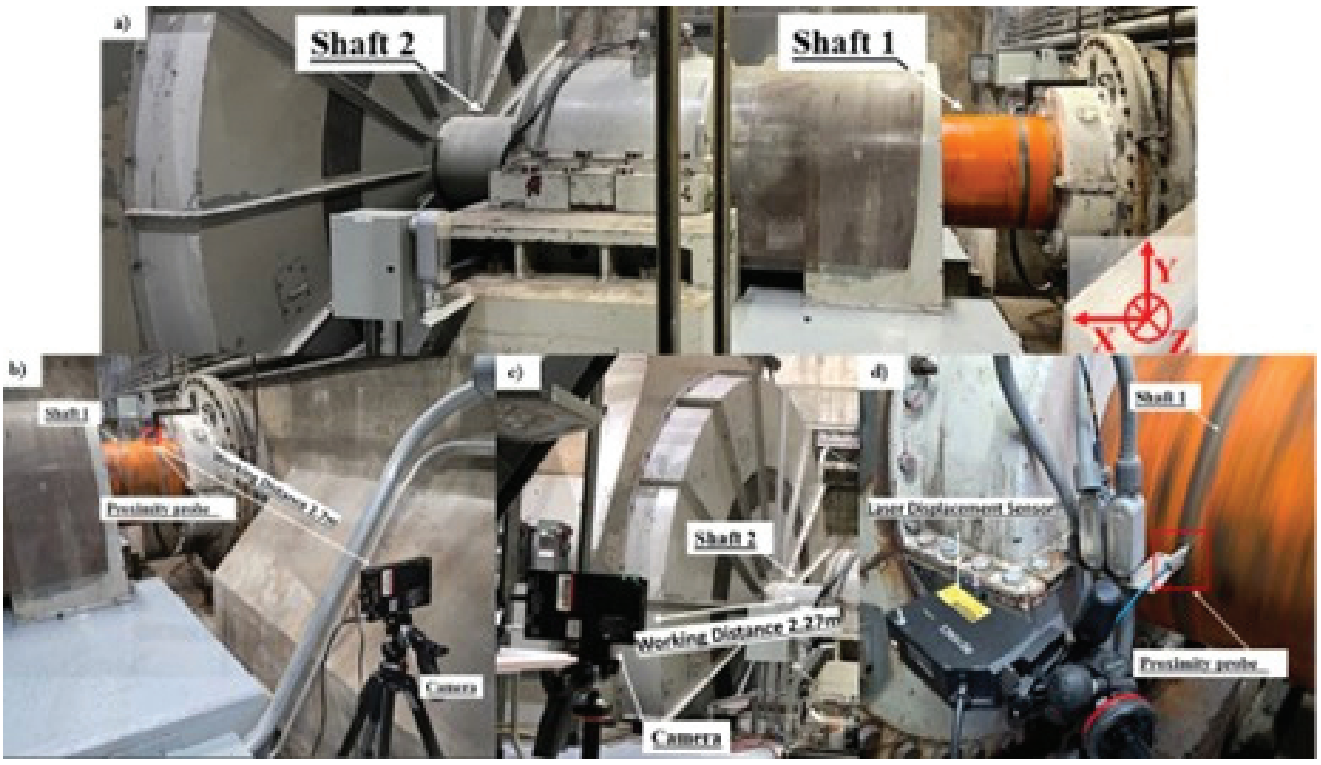


Fig. 6 8.65 MW horizontal Francis turbine at the Boott Hydropower Plant in Lowell, MA used to perform the field tests: a) overview of the turbine showing the two shafts inspected during the tests and indication of the X-Y-Z reference system used, b) experimental setup showing the camera and the proximity probe locations with reference to Shaft 1 used to measure displacements in the Y-direction, c) experimental setup showing the camera locations with reference to Shaft 2 used to measure displacements in the Y-direction, and d) view of the area behind Shaft 1 showing the location of the laser displacement sensor and the second proximity probe to measure displacements in the Z-direction.

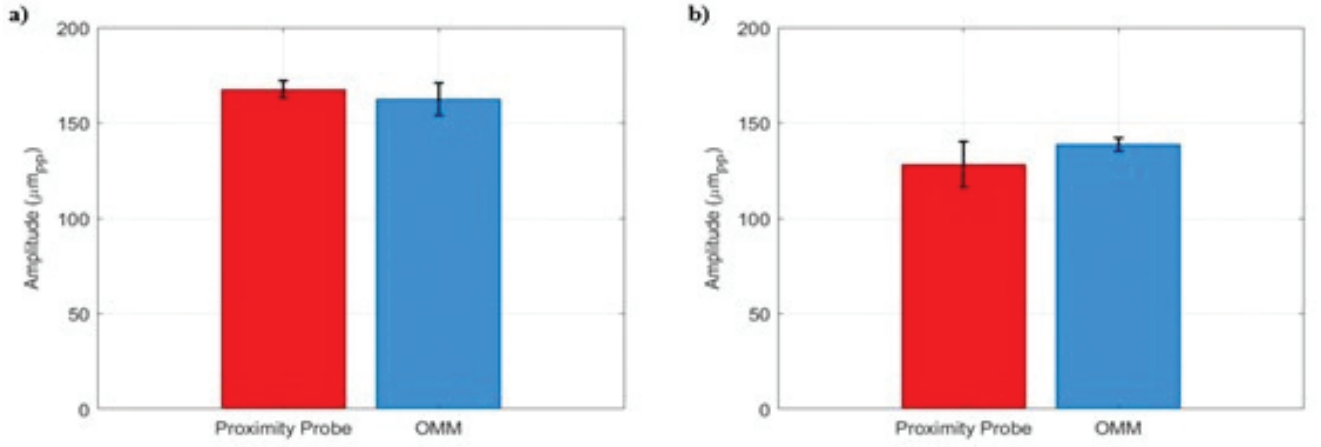


Fig. 7 Comparison of the results obtained with OMM and the proximity probes during the field tests: a) results for shaft 1 and b) results for shaft 2.

Conclusion

In summary, the field tests demonstrated that OMM is capable of capturing small-amplitude displacements on the order of 100 μm in a real hydropower environment, even when the camera is positioned at distances greater than 2 meters from the target. Error margins between OMM and traditional sensors ranged from 3% to 19%, with the overall performance of the

camera-based system deemed reliable for operational monitoring. These findings support the feasibility of deploying OMM for condition monitoring in hydropower applications, though further testing is recommended to refine its capabilities and ensure robustness across varying operational scenarios.

Acknowledgments This work relates to the Department of Navy award N00014-23-1-2124 issued by the Office of Naval Research (ONR). The U.S. Government has a royalty-free license throughout the world in all copyrightable material contained herein. The authors would like to thank the staff and personnel of the Boott Hydropower Plant for providing access to the facility used during the field campaign and assisting with testing.

References

1. Bandara, S., Rajeev, P., & Gad, E. (2024). A review on condition assessment technologies for power distribution network infrastructure. *Structure and Infrastructure Engineering*, *20*(11), 1834-1851.
2. Wang, J., Xu, C., Zhang, J., & Zhong, R. (2022). Big data analytics for intelligent manufacturing systems: A review. *Journal of Manufacturing Systems*, *62*, 738-752.
3. Gagliardi, V., Tosti, F., Bianchini Ciampoli, L., Battagliere, M. L., D'Amato, L., Alani, A. M., & Benedetto, A. (2023). Satellite remote sensing and non-destructive testing methods for transport infrastructure monitoring: Advances, challenges and perspectives. *Remote Sensing*, *15*(2), 418.
4. Zhang, L., Pang, M., Bahaj, A. S., Yang, Y., & Wang, C. (2021). Small hydropower development in China: Growing challenges and transition strategy. *Renewable and Sustainable Energy Reviews*, *137*, 110653.
5. Betti, A., Crisostomi, E., Paolinelli, G., Piazzi, A., Ruffini, F., & Tucci, M. (2021). Condition monitoring and predictive maintenance methodologies for hydropower plants equipment. *Renewable Energy*, *171*, 246-253.
6. Baron, P., Kočiško, M., Hlavatá, S., & Franas, E. (2022). Vibrodiagnostics as a predictive maintenance tool in the operation of turbo generators of a small hydropower plant. *Advances in mechanical engineering*, *14*(5), 16878132221101023.
7. Zhang, G., Yu, R., Dai, L., & Pan, J. (2019). Condition monitoring and fault diagnosis of hydropower station units. *Academic Journal of Engineering and Technology Science*, *2*(2), 89-95.
8. Wadhwa, N., Rubinstein, M., Durand, F., & Freeman, W. T. (2013). Phase-based video motion processing. *ACM Transactions on Graphics (ToG)*, *32*(4), 1-10.
9. Chen, J. G., Wadhwa, N., Cha, Y. J., Durand, F., Freeman, W. T., & Buyukozturk, O. (2015). Modal identification of simple structures with high-speed video using motion magnification. *Journal of Sound and Vibration*, *345*, 58-71.
10. Tiboni, M., Remino, C., Bussola, R., & Amici, C. (2022). A review on vibration-based condition monitoring of rotating machinery. *Applied Sciences*, *12*(3), 972.
11. Romanssini, M., de Aguirre, P. C. C., Compassi-Severo, L., & Girardi, A. G. (2023). A review on vibration monitoring techniques for predictive maintenance of rotating machinery. *Eng*, *4*(3), 1797-1817.
12. Goyal, D., Mongia, C., & Sehgal, S. (2021). Applications of digital signal processing in monitoring machining processes and rotary components: A review. *IEEE Sensors Journal*, *21*(7), 8780-8804.
13. Verellen, T., Verbelen, F., Stockman, K., & Steckel, J. (2021). Beamforming applied to ultrasound analysis in detection of bearing defects. *Sensors*, *21*(20), 6803
14. Dao, F., Zeng, Y., Zou, Y., Li, X., & Qian, J. (2021). Acoustic vibration approach for detecting faults in hydroelectric units: A review. *Energies*, *14*(23), 7840.
15. Kundu, P. (2024). Review of rotating machinery elements condition monitoring using acoustic emission signal. *Expert Systems with Applications*, *252*, 124169.
16. Wang, R., Zhan, X., Bai, H., Dong, E., Cheng, Z., & Jia, X. (2022). A review of fault diagnosis methods for rotating machinery using infrared thermography. *Micromachines*, *13*(10), 1644.
17. Mehta, A., Goyal, D., Choudhary, A., Pabla, B. S., & Belghith, S. (2021). Machine Learning-Based Fault Diagnosis of Self-Aligning Bearings for Rotating Machinery Using Infrared Thermography. *Mathematical Problems in Engineering*, *2021*(1), 9947300.
18. Karna, A., Chitrakar, S., Neopane, H. P., & Dahlhaug, O. G. (2025). A review of condition monitoring in Francis turbines for predictive maintenance. *Journal of Vibration and Control*, 10775463251315816.
19. de Santis, R. B., Gontijo, T. S., & Costa, M. A. (2022). Condition-based maintenance in hydroelectric plants: A systematic literature review. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, *236*(5), 631-646.
20. Niezrecki, C., Baqersad, J., & Sabato, A. (2019). Digital image correlation techniques for NDE and SHM. In *Handbook of advanced nondestructive evaluation* (pp. 1545-1590). Springer, Cham.
21. Sabato, A., Dabetwar, S., Kulkarni, N. N., & Fortino, G. (2023). Noncontact sensing techniques for AI-aided structural health monitoring: a systematic review. *IEEE Sensors Journal*, *23*(5), 4672-4684.
22. Civera, M., Zanotti Fragonara, L., & Surace, C. (2020). An experimental study of the feasibility of phase-based video magnification for damage detection and localisation in operational deflection shapes. *Strain*, *56*(1), e12336.
23. Le Ngo, A. C., Johnston, A., Phan, R. C. W., & See, J. (2018, May). Micro-expression motion magnification: Global lagrangian vs. local eulerian approaches. In *2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018)* (pp. 650-656). IEEE.
24. Valente, N. A., do Cabo, C. T., Mao, Z., & Niezrecki, C. (2022). Quantification of phase-based magnified motion using image enhancement and optical flow techniques. *Measurement*, *189*, 110508.
25. Dragonetti, P., Civera, M., Miraglia, G., & Ceravolo, R. (2025). A Benchmark Dataset for the Validation of Phase-Based Motion Magnification-Based Experimental Modal Analysis. *Data (2306-5729)*, *10*(4).
26. Siringoringo, D. M., Wangchuk, S., & Fujino, Y. (2021). Noncontact operational modal analysis of light poles by vision-based motion-magnification method. *Engineering Structures*, *244*, 112728.

27. Rohe, D. P., & Reu, P. L. (2021). Experimental modal analysis using phase quantities from phase-based motion processing and motion magnification. *Experimental Techniques*, *45*(5), 297-312.
28. Luo, K., Kong, X., Li, J., Hu, J., & Deng, L. (2024). Motion magnification for video-based vibration measurement of civil structures: A review. *Mechanical Systems and Signal Processing*, *220*, 111681.
29. Čufar, K., Slavič, J., & Boltežar, M. (2024). Mode-shape magnification in high-speed camera measurements. *Mechanical Systems and Signal Processing*, *213*, 111336.
30. Civera, M., Zanotti Fragonara, L., Antonaci, P., Anglani, G., & Surace, C. (2021). An Experimental Validation of Phase-Based Motion Magnification for Structures with Developing Cracks and Time-Varying Configurations. *Shock and Vibration*, *2021*(1), 5518163.
31. Cosco, F., Cuenca, J., Desmet, W., Janssens, K., & Mundo, D. (2022). Towards phase-based defect detection: A feasibility study in vibrating panels. *Journal of Sound and Vibration*, *537*, 117196.
32. Nieduzak, T., Valente, N. A., Niezrecki, C., & Sabato, A. (2023, June). Optical Motion Magnification: A Comparative Study and Application for Vibration Analysis. In *Society for Experimental Mechanics Annual Conference and Exposition* (pp. 1-7). Cham: Springer Nature Switzerland.
33. Nieduzak, T. B., Kulkarni, N. N., Niezrecki, C., & Sabato, A. (2023). Wind Turbine Monitoring Using Optical Motion Magnification: Challenges and Opportunities. *STRUCTURAL HEALTH MONITORING 2023*.

Chapter 10
Development of a Novel Robotic Continuously Scanning Laser Doppler Vibrometer for Vibration Measurement of a Beam

Ke Yuan and Weidong Zhu

Abstract A continuously scanning laser Doppler vibrometer (CSLDV) has been widely used to measure structural vibrations due to its high measurement efficiency and spatial resolution. However, a conventional CSLDV, which conducts continuous scanning through two rotating scan mirrors, is limited by its field of view. This work proposes a novel robotic CSLDV system that combines a robotic arm and a laser vibrometer to provide a more flexible approach to obtaining modal parameters of a structure. A central control and data acquisition system is developed to synchronously control the robotic arm, laser vibrometer, and excitation device during the measurement. A new signal processing method is developed to separate vibration signals of the structure from motion signals of the laser vibrometer in the measured response, and to extract mode shapes of the structure. The proposed system and signal processing method are experimentally validated by identifying the first mode shape of a beam under sinusoidal excitation. Mode shapes measured using the conventional CSLDV and the proposed system in a point-by-point manner are used as references. Modal Assurance Criterion values among mode shapes from the three experiments are greater than 99%.

Keywords Laser Doppler vibrometry · CSLDV · robot-assisted vibration measurement · modal identification· signal processing

Introduction

As an optical-based device, the continuously scanning laser Doppler vibrometer (CSLDV) has a limited field of view (FOV) and cannot measure vibrations in areas outside of it. Various studies have been conducted to develop advanced measurement methodologies to extend the FOV of the laser vibrometer. Two common methods are moving the laser vibrometer to different positions [1] and using a reflective mirror to redirect the laser beam of the laser vibrometer [2, 3, 4]. The first method involves building a global coordinate system using reference objects and reconstructing vibrations of different areas based on that system. However, measurement errors can be introduced during the repositioning of the device, and the method becomes less feasible as the number of measurement surfaces increases due to strict experimental setup requirements. The second method keeps positions of both the laser vibrometer and the test structure unchanged to minimize movement errors, redirecting the laser beam to otherwise unmeasurable areas via reflective mirrors. Yuan and Zhu [3] developed a novel mirror-assisted testing methodology, including a new algorithm to design continuous and synchronous scan trajectories using mirrored areas, and a new signal processing method that stitches vibration signals from different parts of the structure using mirrored laser heads for three-dimensional (3D) CSLDV measurements. Although this method supports measurements of structures with complex surfaces, such as cylindrical ones, it requires mirrors of appropriate sizes and carefully placed orientations, which can significantly increase experimental duration.

Ke Yuan
Michael W. Hall School of Mechanical Engineering Mississippi State University, Mississippi State, MS 39762
e-mail: ky309@msstate.edu

Weidong Zhu
Department of Mechanical Engineering
University of Maryland Baltimore County, Baltimore, MD 21250
e-mail: wzhu@umbc.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

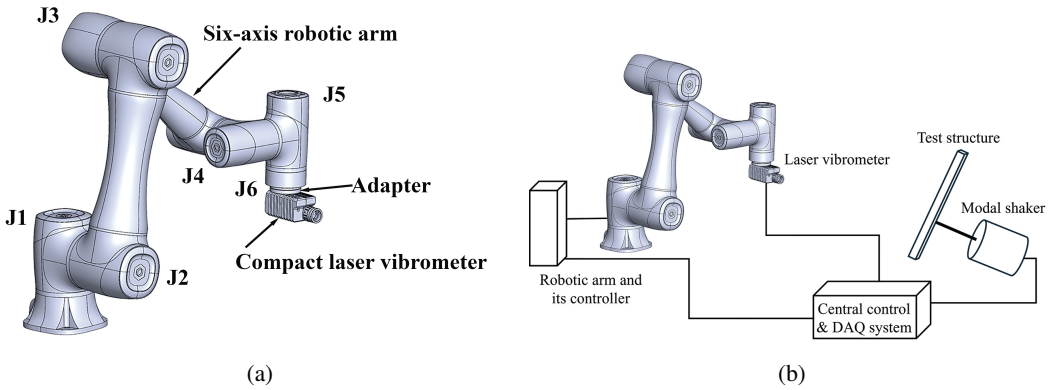
Using robotic platforms such as unmanned aerial vehicles (UAVs) or robotic arms to assist structural health monitoring has received growing attention due to the increased flexibility and accessibility they provide. For example, UAV-assisted crack detection methods have been developed using image processing techniques for bridges with ongoing traffic [5]. UAV-based digital image correlation systems [6, 7] have also been developed to capture dynamic responses of large infrastructures. Mounting a laser vibrometer on a programmable moving platform like a UAV or robotic arm offers a promising approach to overcoming limitations of traditional vibration measurement systems. By inputting programmed positions and orientations into the UAV or the robotic arm, the laser spot can be directed accurately to desired positions on the structure, avoiding manual movement errors. Refs. [8, 9] proposed UAV+LDV solutions and developed correction methods to compensate for angular and linear movement errors. Other studies [10, 11, 12, 13, 14, 15] demonstrated the application of robot+LDV systems for vibration measurements in industrial scenarios, such as damage detection in composite plates and model updating of bladed wheels. In these studies, large industrial robotic arms were often used to accommodate multiple laser heads for full-field 3D vibration measurements, which increased space and cost requirements. In Ref. [16], a 3D robot+LDV system using only one laser head was proposed to reduce the payload requirement of the robotic arm to 3 kg, making the system more compact and suitable for laboratory environments. However, existing robot-assisted LDV systems have primarily focused on step-by-step measurements, and there is a lack of studies on integrating robotic arms with the CSLDV technique to significantly increase measurement efficiency. For instance, authors of Ref. [16] reported that their approach required approximately 20 hours to test 301 measurement points, since additional time was needed for the robotic arm to reposition sequentially among target points. Moreover, the contribution of the robotic motion itself was neglected in their analysis, which may limit both measurement efficiency and accuracy.

The first contribution of this work is the development of a novel robotic CSLDV system for structural vibration measurement. The proposed system consists of a six-axis robotic arm and a compact laser vibrometer. The major difference between the proposed robotic CSLDV system and the conventional CSLDV lies in the scanning mechanism: in the conventional CSLDV, the laser head is stationary and the laser spot moves via rotating scan mirrors, whereas in the robotic CSLDV system, the laser head and the laser spot move together through joint rotations of the robotic arm. Previous research [17, 18, 19] has shown that speckle noise is a major factor affecting the quality of vibration measurements using moving laser vibrometers. In this work, a chip-based laser vibrometer with real-time speckle compensation [20] is used to obtain vibration responses with reduced noise and improved quality. The second contribution of this work is the development of a novel notch-filter-enhanced demodulation (NFED) method to identify modal parameters of structures under sinusoidal excitation. The conventional demodulation method for conventional CSLDV measurements has limited accuracy when applied to signals from the proposed robotic CSLDV system, due to low-frequency interference introduced by the motion of the laser head. The proposed NFED method can separate vibration signals of a structure from motion signals of the laser vibrometer in the measured response and extract mode shapes of the structure. The proposed system and signal processing method are experimentally validated by identifying the first mode shape of a beam under sinusoidal excitation. Mode shapes measured using the conventional CSLDV and the proposed system in a point-by-point manner are used as references. Modal Assurance Criterion (MAC) values among mode shapes from the three experiments are greater than 99%. In addition, the effect of the running speed of the proposed system on the estimated mode shape is investigated in this work. Compared to previous robot+LDV studies, the proposed robotic CSLDV system leverages continuous programmed motion of the robotic arm to achieve vibration measurements along arbitrary scan trajectories in a fraction of the time, while explicitly addressing robot-induced motion effects through the proposed NFED signal processing method.

Methodology

As shown in Figure 1a, the proposed system consists of a compact laser vibrometer and a six-axis robotic arm, which are connected via an adapter. Six joints of the robotic arm, from its base to its tip, are denoted as J1-J6, and and each can be independently controlled by an internal motor. During vibration measurement, the movement of the robotic arm, DAQ of the laser vibrometer, and excitation provided by a modal shaker are synchronized by a central control and DAQ system, as shown in Figure 1b.

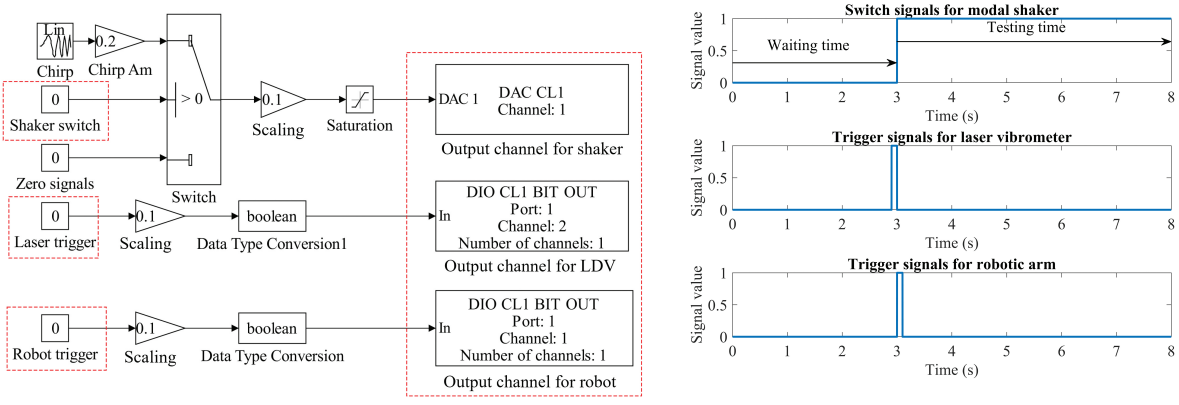
The working principle of the central control and DAQ system via Simulink is illustrated in Figure 2a. The control model consists of three tracks corresponding to the modal shaker, the laser vibrometer, and the robotic arm, respectively. Initial signals for the three blocks denoting shaker switch, laser trigger, and robot trigger, respectively, are all set to zero, and digital signals are applied to initiate vibration measurements. The gain blocks of 0.1 are used as required scalings in the model because a value of 1 V in a digital-to-analog converter (DAC) corresponds to an output of 10 V. Figure 2b provides an example of digital signals used during testing. In the top track, the modal shaker is controlled by a switch block. Its data inputs, shown in the first and third ports, are a series of excitation and zero signals, respectively. The control input, shown in



(a) (b)

Fig. 1 (a) Proposed novel robotic CSLDV system, consisting of a six-axis robotic arm and a compact laser vibrometer, and (b) the working principle of the robotic CSLDV system, where the robotic arm, the laser vibrometer, and the modal shaker are synchronized by a central control and DAQ system

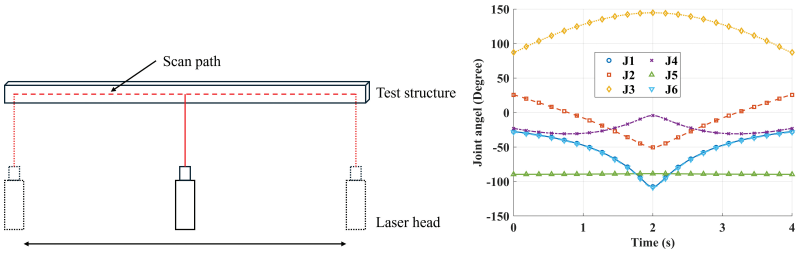
the second port, has a threshold value of zero. Data are sent to the modal shaker via the DAC, shown at the right end of the top track. The saturation block in the top track is used for limiting the range of signals to avoid undesired sensor overload. From Figure 2b, the control input consists of two phases: a waiting time and a testing time. During the waiting time, the control input signal remains zero, which does not meet the threshold; therefore, the switch block passes zero signals, and the modal shaker remains disabled. At the end of the waiting time, which is 3 seconds in this example, the signal value switches to one, meeting the threshold. As a result, the switch block passes excitation signals, which are chirp signals in this example, and the modal shaker is activated. The middle and bottom tracks in Figure 2a show triggers for DAQ of the laser vibrometer and the movement of the robotic arm, respectively. Both tracks follow similar logic: a digital pulse is used as the trigger signal near the start of the testing time. The only difference is that the laser vibrometer is triggered by the falling edge of the pulse, while the robotic arm is triggered by the rising edge.



(a) (b)

Fig. 2 (a) Working principle of the central control and DAQ system, and (b) switch and trigger signals for the modal shaker, the laser vibrometer, and the robotic arm, used to synchronously control all three components

Continuous scanning measurements can be performed using the proposed system by continuously moving the laser vibrometer. This approach is significantly different from a conventional CSLDV system, which performs continuous scanning by rotating a pair of orthogonal scan mirrors. An example of using the proposed robotic CSLDV system to scan a beam-like structure is shown in Figure 3a, where the laser vibrometer is translated along a straight line on the surface of the beam at a constant speed. A complete scan in this study is defined as the laser spot moving from the start point to the end point and then returning. The period of one scan is denoted as T, and the scan frequency is given by $f_{sc} = 1/T$. The distance between the laser head and the surface of the test beam remains constant during scanning, indicating that the tool center point (TCP) of the system follows a linear trajectory.



(a) (b)

Fig. 3 (a) An example of a scan path executed by the proposed system along a straight line on the surface of a beam, and (b) time series of rotation angles of robotic arm joints J1-J6 corresponding to the straight-line path

Kinematics of the proposed robotic CSLDV system can be described using the Denavit–Hartenberg (DH) convention [21],

$$ {}^{0}_{la}T = T_1^0 T_2^1 T_3^2 T_4^3 T_5^4 T_6^5 T_{la}^6 \tag{1} $$

where ${}^{0}_{la}T$ denotes the transformation matrix describing the pose of the laser spot, which is an end-effector in this work, relative to the base of the system, T_i^{i-1} denotes the relative pose of the ith joint to the $(i-1)^{\text{th}}$ joint, and T_{la}^6 represents the pose of the laser spot relative to joint J6. Each T_i^{i-1} is determined by DH parameters, including link lengths a, link offsets d, link twists α, and joint angles θ. The first three parameters are known from the robotic arm's physical configuration. T_{la}^6 can be determined from the laser head geometry and its distance to the test structure. Inverse kinematics (IK) based on DH parameters of the system is used to calculate its joint angles θ as

$$ \theta_i = f^{-1}({}^{0}_{la}P) \tag{2} $$

where ${}^{0}_{la}P$ denotes the position of the laser spot in the base coordinate system.

In this work, a commercial Collaborative robot (Cobot) model DOBOT CR5A (Shenzhen Dobot Corp Ltd) and its control software, DobotStudio, are used to solve the IK equations and obtain joint angles. As a Cobot, the CR5A allows teaching by manually guiding the end-effector to the desired start and end points of the scan path in drag mode. These positions are recorded and used as input parameters for the IK. The joint rotation angles for a straight-line path at $f_{sc} = 0.25$ Hz are shown in Figure 3b. The joint J6 exhibits the largest rotation of 80.1 degrees from start to end, while the joint J5 rotates by only 1 degree. To maintain a constant TCP speed, each joint rotates at a different rate.

The response of the beam without excitation, as shown in Figure 4a, is measured using the displacement-based laser vibrometer in the proposed system to verify the scan path. The measured response has an amplitude of less than 0.8 mm, as indicated by the dashed lines, suggesting that the distance between the start and end points is also less than 0.8 mm. Given the beam length of 548 mm, the angle error between the actual and designed scan paths is approximately 0.08 degrees, indicating high accuracy in laser spot positioning. The response of the beam under sinusoidal excitation at 66.67 Hz, which is close to the first natural frequency of the beam, is shown in Figure 4b. The system captures both the low-frequency motion of the laser head and the high-frequency vibration of the beam, demonstrating an advantage of the displacement-based laser vibrometer over conventional velocity-based models.

In conventional CSLDV measurements, the output response of a structure is a vibration signal modulated by spatial deflection due to continuous scanning motion [22]. When a straight-line scan trajectory is used to direct the laser spot across the structure, the response distribution $u(x,t)$ can be expressed as

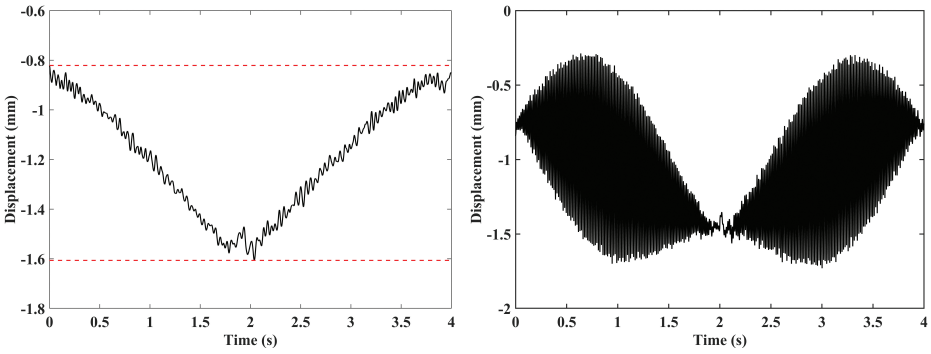
$$ u(x,t) = U_I(x)\cos(\omega_e t) + U_Q(x)\sin(\omega_e t) \tag{3} $$

where $U_I(x)$ and $U_Q(x)$ represent in-phase (real) and the quadrature (imaginary) components of the mode shape of the structure, ω_e is the excitation frequency, and x denotes the spatial coordinate along the scan path, normalized within [-1,1]. For a CSLDV measurement with a uniform scanning rate, in which $x = f(t)$, this spatially varying response can be expressed purely as a function of time [23]:

$$ u(t) = u_I(t)\cos\omega_e t + u_Q(t)\sin\omega_e t \tag{4} $$

To extract u_I and u_Q, the response is multiplied by cosine and sine functions of the excitation frequency, yielding:

$$ \begin{aligned} u(t)\cos\omega_e t &= u_I(t)\cos^2\omega_e t + u_Q(t)\sin\omega_e t\cos\omega_e t \\ &= \frac{1}{2}u_I(t) + \frac{1}{2}u_I(t)\cos 2\omega_e t + \frac{1}{2}u_Q(t)\sin 2\omega_e t \end{aligned} \tag{5} $$



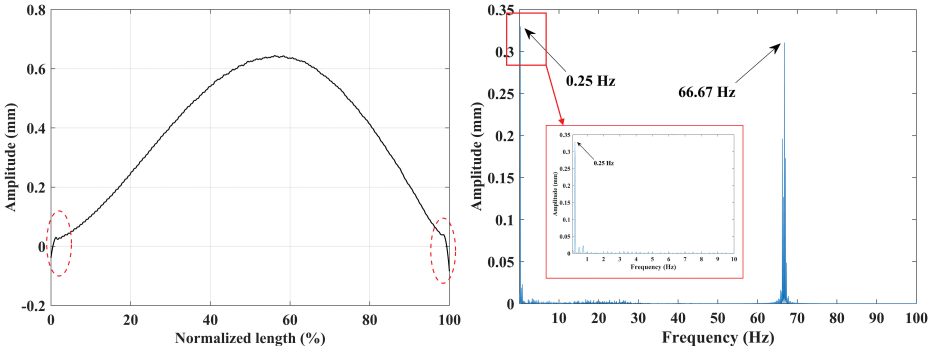
(a) (b)

Fig. 4 Measured responses of the beam: (a) without excitation, and (b) under sinusoidal excitation at 66.67 Hz

$$\begin{aligned} u(t)\sin\omega_e t &= u_I(t)\cos\omega_e t\sin\omega_e t + u_Q(t)\sin^2\omega_e t \\ &= \frac{1}{2}u_Q(t) + \frac{1}{2}u_I(t)\sin 2\omega_e t - \frac{1}{2}u_Q(t)\cos 2\omega_e t \end{aligned} \tag{6}$$

Signal components with frequency $2\omega_e$ can be removed by applying a low-pass filter, and scaling the filtered signals by a factor of 2 yields the desired components $u_I(t)$ and $u_Q(t)$. Finally, mode shape components can be obtained by substituting $t = f^{-1}(x)$ into $u_I(t)$ and $u_Q(t)$.

This method is referred to as demodulation, as it separates the vibration signal from spatial modulation effects. To test the applicability of the demodulation method to the proposed robotic CSLDV system, it was used to process the response shown in Figure 4b. As seen in Figure 5a, the resulting first mode shape of the beam is not smooth and exhibits noticeable boundary distortions, as highlighted by dashed circles, differing significantly from results obtained using the conventional CSLDV [24]. To investigate the cause of this inaccuracy, a frequency spectrum of the measured response was obtained using a fast Fourier transform (FFT), as shown in Figure 5b. The spectrum displays two dominant peaks at 0.25 Hz and 66.67 Hz, corresponding to the laser head motion and the beam vibration, respectively. The proximity of the low-frequency motion signal to zero frequency reduces the effectiveness of the low-pass filter in the demodulation method, resulting in inaccurate mode shape identification.



(a) (b)

Fig. 5 (a) First mode shape of the beam identified from the response shown in Figure 4b using the conventional demodulation method, and (b) the frequency spectrum of the measured response, displaying two dominant peaks at 0.25 Hz and 66.67 Hz

To overcome this limitation, a notch-filter-enhanced demodulation (NFED) method is proposed. In this method, notch filters are applied to the measured response to isolate the low-frequency laser head motion and high-frequency beam vibration. Each notch is implemented as a second-order infinite impulse response filter with complex-conjugate zeros on the unit circle at the target frequency and slightly retracted poles to set the notch selectivity. The discrete-time transfer function is

$$H(z) = \frac{1 - 2\cos(\omega_0)z^{-1} + z^{-2}}{1 - 2r\cos(\omega_0)z^{-1} + r^2 z^{-2}} \tag{7}$$

where $\omega_0 = 2\pi F_0/F_s$, F_s is the sampling rate, F_0 is the center frequency of the notch that can be set to either the scan frequency f_{sc} or the excitation frequency f_e, and $r \in (0, 1)$ is the pole radius that controls the notch bandwidth. In this work, $r = 0.99$ is used to achieve a narrow, selective notch around F_0.

To isolate the structural vibration for demodulation, the measured signal is filtered with a notch centered at f_{sc} by removing laser head motion. Conversely, to visualize the robot-motion component, a notch centered at f_e can be applied. Separated signals are shown in Figure 6a, with the top subplot showing the beam vibration and the bottom subplot showing the laser head motion. Their respective frequency spectra are provided in Figure 6b, each exhibiting a single prominent peak. With the beam vibration signal filtered, its mode shape can be accurately extracted using the demodulation method, as shown in Figure 6c.This enhancement addresses limitations of the conventional demodulation approach for the proposed robotic CSLDV system.

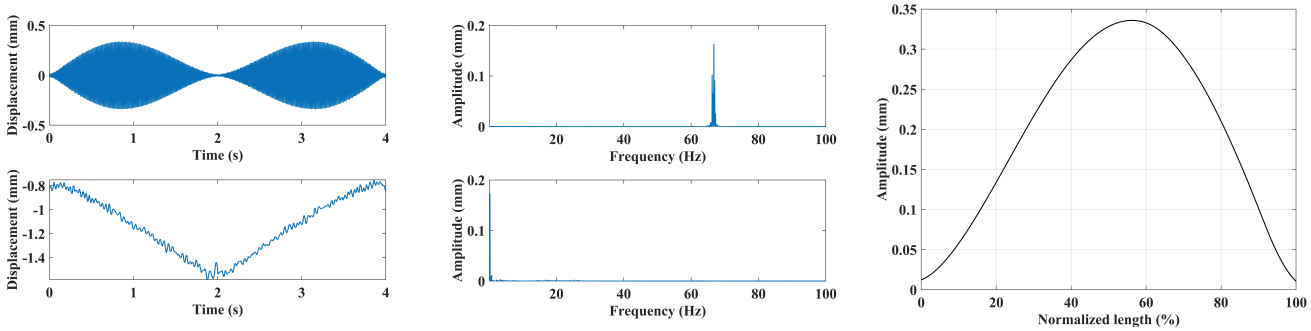


Fig. 6 Separated signals using notch filters applied to the measured response of the beam: (a) time domain and (b) frequency domain, where the top subplots show the beam vibration and the bottom subplots show the laser head motion, and (c) the first mode shape of the beam identified using the proposed NFED method

Experimental investigation

In this section, a series of experiments were conducted to validate the proposed robotic CSLDV system for modal identification of a beam under sinusoidal excitation. The experimental setup is shown in Figure 7, where Figure 7a displays the initial position of the robotic CSLDV system, and Figure 7b shows the test position. The six-axis robotic arm used in this study is the DOBOT CR5A model (Shenzhen Dobot Corp. Ltd., Shenzhen, China), and the laser vibrometer is the MotionGo model (OmniSensing Photonics LLC, Maryland, USA). A machined flange was used to connect the vibrometer to the robotic arm. The payload capacity of the CR5A is 5,000 g, while the total weight of the MotionGo vibrometer and the flange is 530 g that is well within the payload limit of the robotic arm.

In the system's initial position, where the rotation angles of all six joints are set to zero, the laser beam is parallel to the beam surface, meaning that the test beam lies outside the FOV of the system. This configuration is difficult to measure using the conventional CSLDV. The test beam is clamped at one end in a bench vise and excited at the other end using a modal shaker driven by a sinusoidal signal with a frequency of 66.67 Hz, which is close to its first natural frequency. In the test position, the laser head moves along the scan direction shown in Figure 7b, which is parallel to the beam surface. The total scan length is 548 mm, and the joint angles for this scanning motion are identical to those shown in Figure 3.

In this experiment, the running speed of the system was set to 274 mm/s, corresponding to a scan frequency of 0.25 Hz. Five complete scans, taking a total of 20 seconds, were conducted. Vibration signals of the beam, separated from the measured response using the NFED method, are shown in Figure 8a. An averaging technique was applied to the vibration signals to obtain the response from one complete scan, as shown in Figure 8b. This averaged signal was used to extract the mode shape of the beam using the proposed NFED method. In addition, two reference experiments were conducted to validate the mode shape identified by the robotic CSLDV system. The first reference experiment that can be referred to as reference 1 used the conventional CSLDV with the same scan frequency of 0.25 Hz. The second reference experiment that can be referred reference 2 used the proposed robotic CSLDV system in a point-by-point scanning mode, where the system paused at each point for measurement before moving to the next one.

Figure 8c compares the first mode shapes of the beam obtained from all three experiments, where the solid line denotes the first mode shape of the beam identified using the proposed robotic CSLDV system, the dashed line denotes that identified

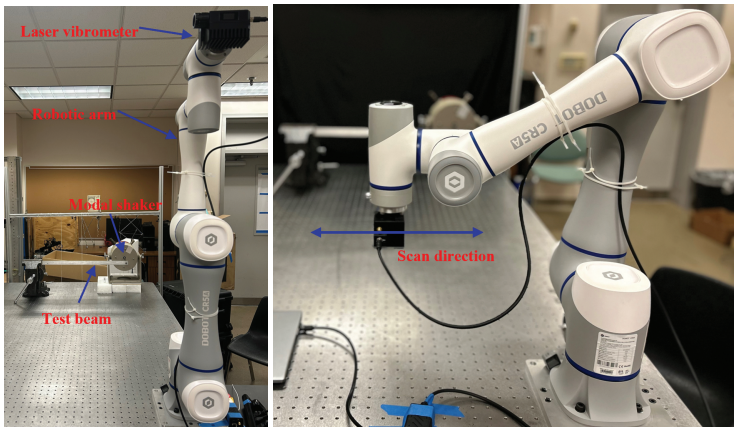


Fig. 7 (a) Experimental setup of vibration measurement of the beam using the robotic CSLDV system in its initial position, where the laser beam is parallel to the beam surface, and (b) the test position of the robotic CSLDV system, where the arrow indicates its scan direction

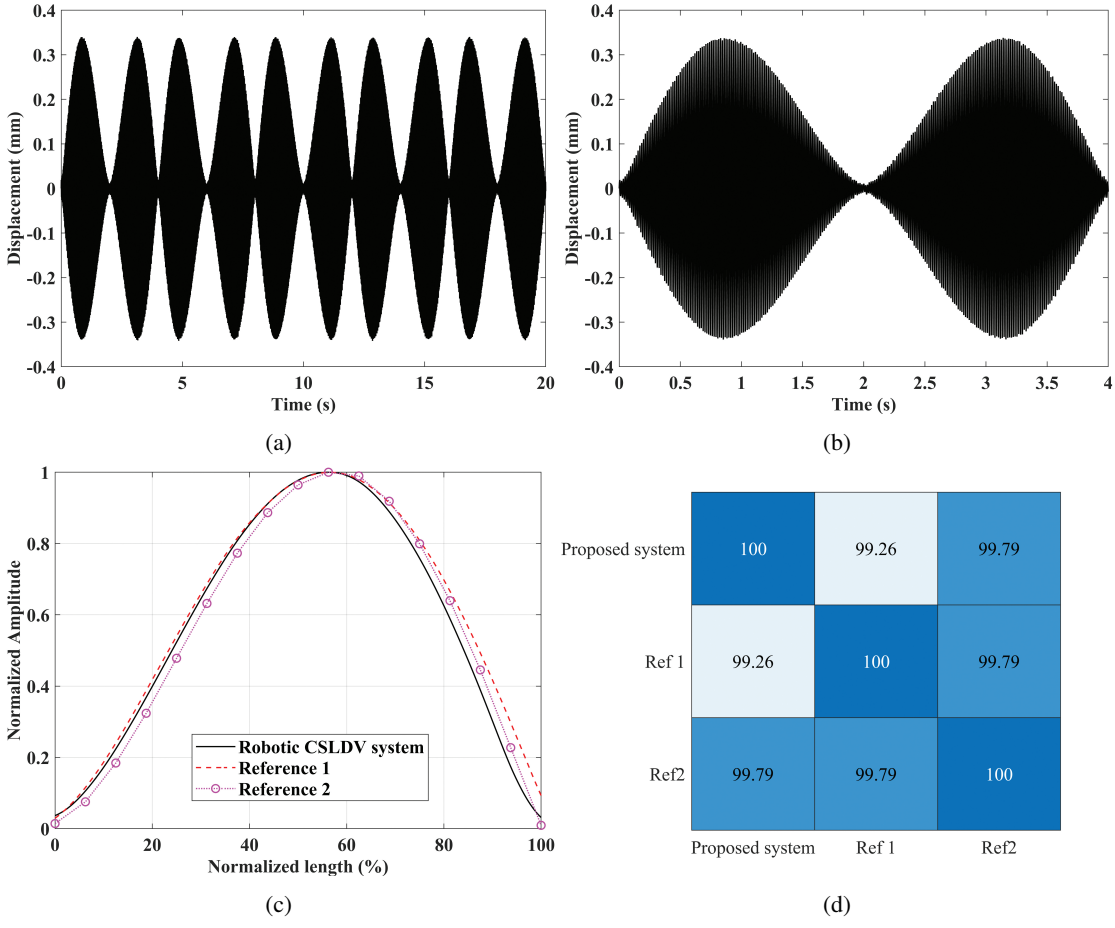


Fig. 8 (a) Vibration signals of the beam measured over five complete scans, (b) averaged vibration signal from one complete scan, used to extract the mode shape of the beam using the proposed NFED method, (c) first mode shapes of the beam identified using the proposed robotic CSLDV system, the conventional CSLDV, and the point-by-point method, and (d) MAC values among the identified mode shapes

using the conventional CSLDV, and the line with circular markers denotes that identified in the point-by-point way. The mode shape identified by the proposed system closely matches those from both reference experiments. To quantitatively evaluate the correlation among the identified mode shapes, MAC values [25] were calculated using:

$$\mathrm{MAC}(\{\varphi_r\},\{\varphi_s\}) = \frac{|\{\varphi_r\}^{\mathrm{T}}\{\varphi_s\}|^2}{(\{\varphi_r\}^{\mathrm{T}}\{\varphi_r\})(\{\varphi_s\}^{\mathrm{T}}\{\varphi_s\})} \times 100\% \tag{8}$$

where φ_r and φ_s are modal vectors of the rth and s^{th} modes of the structure, respectively, and the superscript T denotes matrix transpose. A MAC value close to 100% indicates strong correlation between two mode shapes, while a value near 0 indicates weak correlation. The resulting MAC matrix is shown in Figure 8d. The off-diagonal MAC values between the mode shape identified by the proposed system and those from the two reference experiments are both greater than 99%, confirming the high consistency among them. Comparison among test duration, number of measurement points, and availability beyond the FOV of the three measurement systems is shown in Table 1. While all three methods achieve comparable accuracy in identifying the mode shape of the beam, the proposed robotic CSLDV system offers an extended FOV compared to the conventional CSLDV and reduced test duration with higher spatial resolution compared to the point-by-point method.

Table 1 Comparison among test duration, number of measurement points, and availability beyond the FOV of the three measurement systems

System	Test duration (s)	Number of measurement points	Availability beyond the FOV
Robotic CSLDV	20	80,000	Yes
Conventional CSLDV	20	80,000	No
Single-point LDV	101	17	No

Conclusions

A novel robotic CSLDV system, which integrates a six-axis robotic arm and a compact laser vibrometer, was developed in this study to measure structural vibrations with high spatial resolution and scanning flexibility. A NFED method was proposed to separate structural vibration signals from the motion-induced components of the laser vibrometer. Experimental validation and comparative studies were carried out to assess the accuracy and repeatability of the system in identifying beam mode shapes under sinusoidal excitation. Main conclusions from this work are summarized below:

1. A robotic CSLDV system was developed to perform continuous scanning along arbitrary paths via robot arm motion, offering extended FOV compared to conventional CSLDV systems using scan mirrors.
2. A new NFED method was proposed, enabling accurate extraction of structural vibration signals even when robotic motion introduces low-frequency noise near zero frequency in the measurement signal.
3. Experimental validation confirmed high accuracy of the proposed system: the identified first mode shape of a cantilever beam matched well with reference measurements, with MAC values exceeding 99% compared to conventional CSLDV and point-by-point methods.

Acknowledgments This research was supported by the National Science Foundation through Grant No. CMMI-1763024 and the Maryland Technology Development Corporation through the Maryland Innovation Initiative Grant.

References

1. Chen, Y., Escalera Mendoza, A., and Griffith, D. "Experimental dynamic characterization of both surfaces of structures using 3d scanning laser Doppler vibrometer". *Experimental Techniques*, 47(5):989–1006 (2023).
2. Witt, B., Rohe, D., and Schoenherr, T. "Full-field strain shape estimation from 3D SLDV". In *Rotating Machinery, Optical Methods & Scanning LDV Methods, Volume 6: Proceedings of the 37th IMAC, A Conference and Exposition on Structural Dynamics 2019*, pages 31–45. Springer (2019).

3. Yuan, K. and Zhu, W. "A novel mirror-assisted method for full-field vibration measurement of a hollow cylinder using a three-dimensional continuously scanning laser Doppler vibrometer system". *Mechanical Systems and Signal Processing*, 216:111428 (2024).
4. Yuan, K., Yuan, S., and Zhu, W. "Full-field modal analysis of a tensegrity column using a three-dimensional scanning laser Doppler vibrometer with a mirror". *Journal of Vibration and Acoustics*, 146(6):061101 (2024).
5. Gadiraju, D.S., Azam, S.E., and Khazanchi, D. "Shm-traffic: DRL and transfer learning based UAV control for structural health monitoring of bridges with traffic". *IEEE Access*, 13:6166–6179 (2025).
6. Jeong, E., Seo, J., and Wacker, J. "New bridge inspection approach with joint UAV and DIC system". In *Structures Congress 2020*, pages 349–359. American Society of Civil Engineers Reston, VA (2020).
7. Lavezzi, G., Ciarcià, M., Won, K., and Tazarv, M. "A DIC-UAV based displacement measurement technique for bridge field testing". *Engineering Structures*, 308:117951 (2024).
8. Garg, P., Moreu, F., Ozdagli, A., Taha, M.R., and Mascareñas, D. "Noncontact dynamic displacement measurement of structures using a moving laser Doppler vibrometer". *Journal of Bridge Engineering*, 24(9):04019089 (2019).
9. Garg, P., Nasimi, R., Ozdagli, A., Zhang, S., Mascarenas, D.D.L., Reda Taha, M., and Moreu, F. "Measuring transverse displacements using unmanned aerial systems laser Doppler vibrometer (UAS-LDV): Development and field validation". *Sensors*, 20(21):6051 (2020).
10. Oliver, D.E., Schuessler, M., et al. "Automated robot-based 3D vibration measurement system". *Sound Vib*, 43:12–15 (2009).
11. Ziemer, T. and Rembe, C. "Demonstration of a 3D vibration measurement with a single-point laser Doppler vibrometer and a 6-axis industrial robot". In *Proceedings of the International Symposium on Optoelectronic Technology 2016*, pages 1–10 (2016).
12. Bertini, L., Neri, P., Santus, C., and Guglielmo, A. "Automated experimental modal analysis of bladed wheels with an anthropomorphic robotic station". *Experimental Mechanics*, 57:273–285 (2017).
13. Devigne, O., Hoffait, S., and Bruls, O. "Development of a robot-aided modal analysis measurement method using laser Doppler vibrometry". In *ISMA 2020 (29th Leuven Conference on Noise and Vibration Engineering)*, pages 1847—-1858 (2020).
14. Ahmed, H., Mohsin, A., Hong, S.C., Lee, J.R., and Ihn, J.B. "Robotic laser sensing and laser mirror excitation for pulse-echo scanning inspection of fixed composite structures with non-planar geometries". *Measurement*, 176:109109 (2021).
15. Brethee, K.F., Uwayed, A.N., and Alden Qwam, A.Y. "A novel index for vibration-based damage detection technique in laminated composite plates under forced vibrations: Experimental study". *Structural Health Monitoring*, 22(5):3109–3125 (2023).
16. Margerit, P., Gobin, T., Lebée, A., and Caron, J.F. "The robotized laser Doppler vibrometer: On the use of an industrial robot arm to perform 3D full-field velocity measurements". *Optics and Lasers in Engineering*, 137:106363 (2021).
17. Bendel, K., Fischer, M., and Schuessler, M. "Vibrational analysis of power tools using a novel three-dimensional scanning vibrometer". In *Sixth international conference on vibration measurements by laser techniques: advances and applications*, volume 5503, pages 177–184. SPIE (2004).
18. Sels, S., Vanlanduit, S., Bogaerts, B., and Penne, R. "Three-dimensional full-field vibration measurements using a handheld single-point laser doppler vibrometer". *Mechanical Systems and Signal Processing*, 126:427–438 (2019).
19. Zeng, Y., Nunez, A., and Li, Z. "Speckle noise reduction for structural vibration measurement with laser Doppler vibrometer on moving platform". *Mechanical Systems and Signal Processing*, 178:109196 (2022).
20. Yuan, K., Zhu, Z., Chen, W., and Zhu, W. "Development and validation of a new type of displacement-based miniatured laser vibrometers". *Sensors*, 24(16):5230 (2024).
21. Denavit, J. and Hartenberg, R.S. "A kinematic notation for lower-pair mechanisms based on matrices". *Journal of Applied Mechanics*, 22(2):215–221 (2021).
22. Sriram, P., Hanagud, S., and Craig, J. "Mode shape measurement using a scanning laser Doppler vibrometer". *The International Journal of Analytical and Experimental Modal Analysis*, 7(3):169–178 (1992).
23. Stanbridge, A. and Ewins, D. "Modal testing using a scanning laser Doppler vibrometer". *Mechanical Systems and Signal Processing*, 13(2):255–270 (1999).
24. Chen, D.M., Xu, Y., and Zhu, W. "Experimental investigation of notch-type damage identification with a curvature-based method by using a continuously scanning laser Doppler vibrometer system". *Journal of Nondestructive Evaluation*, 36(2):38 (2017).
25. Ewins, D. *Modal Testing: Theory, Practice and Application*. Engineering dynamics series. Wiley (2000).

Chapter 11
A Brief Review of Computer Vision Technology in Pavement Management

Mohamed Tareq Soliman, Feras Abla, Bilal Al-Oubaidi, Onur Avci, James Bryce, and Mahdi Nasimifar

Abstract The integration of computer vision technologies into pavement management represents a paradigm shift from traditional, manual inspection methods towards automated, efficient, and data-driven approaches. This paper presents a brief review on state-of-the-art computer vision methodologies, starting from the process of obtaining data through diverse data acquisition technologies such as 3D laser scanning, RGB-D sensors, ground-penetrating radar (GPR), and unmanned aerial vehicles (UAVs), reaching to developing advanced deep learning models (integration of YOLO object detectors or R-CNN with transformer-based networks). These technologies offer significant benefits, including high accuracy in object detection and segmentation, real-time processing capabilities, substantial reductions in inspection time and cost, improved worker safety by minimizing roadside exposure, and the generation of quantifiable data for precise maintenance planning. However, they face persistent challenges, including sensitivity to environmental conditions (e.g., poor lighting, shadows), difficulties in generalizing across diverse pavement types and complex defect morphologies, computational intensity, data imbalance in training datasets, and the high cost of some acquisition systems. Despite these hurdles, the continued evolution of computer vision holds immense potential for developing fully automated, robust, and comprehensive pavement condition.

Keywords Pavement Condition · Crack Segmentation · CNN.

Introduction

The application of computer vision technology to pavement management has grown dramatically in recent years, driven by the need for efficient automated inspection systems [1]. Manual visual inspection of pavement cracks is extremely labor-intensive and relies entirely on human judgment [2, 3]. Additionally, manual inspections require obstructing traffic and can be dangerous [4]. Computer vision approaches involve systematic acquisition and analysis of images to identify and classify distress regions [5].

The implementation of computer vision-based distress detection and evaluation systems represents a significant shift toward automated infrastructure evaluation [6]. These systems have become increasingly important for promoting data-driven road maintenance, offering continuous and timely monitoring capabilities that are less labor-intensive and more cost-effective than manual inspection methods [6]. Despite these contributions, computer vision-based systems face significant challenges with similarity in features between cracks and background noise in poor lighting, complex textures, blurred boundaries between tiny cracks and local noise, and low-quality, blurry images collected at high speeds [3].

Typical image processing methods for pavement crack detection primarily include edge detection [7, 8], threshold segmentation [9], and region growing techniques [10, 11]. Edge detection algorithms identify crack boundaries using mathematical operators such as Sobel, Prewitt, and Canny operators [7, 10]. The Canny edge detection operator has been particularly effective when combined with morphological operations for automating crack measurements of length, width, and area [12].

Mohamed Tareq Soliman · Feras Abla · Bilal Al-Oubaidi · Onur Avci · James Bryce
Department of Civil and Environmental Engineering, West Virginia University, Morgantown, WV 26506
e-mail: mts00029@mix.wvu.edu; fa00060@mix.wvu.edu; bia00002@mix.wvu.edu; onur.avci@mail.wvu.edu; james.bryce@mail.wvu.edu

Mahdi Nasimifar
Engineering & Software Consultants LLC, Turner Fairbank Highway Research Center, McLean, VA 22101
e-mail: mnasimifar@agr-us.com

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

On the other hand, threshold segmentation methods separate crack regions from pavement backgrounds by setting appropriate pixel intensity thresholds. These approaches assume that cracks have lower intensity values than the surrounding pavement surface [13]. Region growing methods complement these techniques by generating pixel-level labels and performing segmentation of neighboring pixels with similar properties such as grayscale [14]. Several specialized filtering and analysis methods have been developed for pavement crack detection. Gabor filters have proven effective for analyzing textures with specific orientation and frequency characteristics, making them suitable for identifying cracks of various orientations and types [15–17]. Wavelet transform-based methods have been proposed by adjusting the wave in scale (stretched or shrunk) to correspond with the form of a signal and translated along the entire signal to capture various locations of cracks [18]. Computer vision systems have also incorporated Hough transforms for region-of-interest detection and local binary patterns for texture analysis [19].

Data Acquisition Technologies

The performance of all these image processing techniques is directly tied to the dataset they're trained on. Datasets are usually obtained through these methods:

Three-dimensional (3D) Laser Scanning

3D laser scanning accounts for 95% of the reported efforts for acquiring high-resolution, full-coverage 3D pavement surface data, accounting for over 95% of the reported efforts in 2015 [20]. 3D laser scanners are particularly advantageous for their insensitivity to varying lighting conditions and poor intensity contrast, which often hamper 2D image-based methods [21]. Furthermore, 3D data provides crucial depth information for accurate crack detection and allows for precise measurements of defects like rutting and raveling [7, 21]. Although laser scanning is the predominant technology for pavement distress measurement, a wide range of other 3D imaging techniques (e.g., stereo vision, shape-from-focus, interferometry) offer viable and often superior alternatives for specific applications [20].

Kaartinen et al.'s [22] comprehensive review study demonstrated that Terrestrial Laser Scanning (TLS) and Mobile Laser Scanning (MLS) systems generate essential, high-resolution 3D point clouds for contactless detection and assessment of distresses manifesting in height and depth. However, LiDAR can be extremely costly, computationally intense, and vulnerable to surface properties, scanner position, and weather/light conditions [22].

To address the technical limitations of the technology, Yi et al. [23] proposed a framework integrating LiDAR, deep neural networks, and Building Information Modeling (BIM) to automatically detect and manage distresses. The framework localizes defects, including alligator cracks, a combined crack class, and potholes, by drawing 3D bounding boxes around them. The proposed model 'PointPillar' achieved a 3D mean average precision (3D mAP) of 62.7% and a recall of 78.5%, outperforming other 3D detection models in accuracy and inference time [23].

RGB-D Sensors

Inexpensive consumer-grade RGB-D sensors provide a unique combination of 2D visual and 3D depth data at a fraction of the cost of LiDAR or professional laser scanners. Asadi et al. [24] used an inexpensive RGB-D sensor mounted on a car and a Raspberry Pi 4 to run deep learning models for crack detection. The best-performing model (Faster R-CNN with ResNet 101 and reduced region proposals) achieved 97.6% accuracy and ran at 12 fps. However, the dataset lacked very small crack instances and other pavement distresses (e.g., potholes, rutting).

Huang et al. [25] developed an end-to-end road condition monitoring framework with feature-level fusion of RGB integrated with surface normal (SN) data. The dataset consisted of RGB-D image pairs manually annotated and including potholes and non-defective surfaces such as manholes. The proposed system yielded a mean IoU of 89.8%. The model also showed a substantial performance in detecting potholes even in poor lighting conditions.

Another study by Mahmoudzadeh et al. [26] used an affordable RGB-D sensor to perform fusion of color and depth data. This fusion enabled the researchers to reconstruct 3D pavement surfaces and estimate International Roughness Index (IRI) values. The results, which were validated against rod-and-level measurements on both a rough and a smooth road section, showed a strong correlation ($R^2 > 0.95$) and high accuracy (90%). While these findings are promising given the sensor's low cost, the study's limitations include its use of a static data collection method—which is less practical compared to collecting data at traffic speeds—and the small sample size of only two road sections. Furthermore, the impact of environmental variables was not thoroughly investigated.

Ground-Penetrating Radar (GPR)

The high-frequency electromagnetic waves emitted from ground-penetrating radar (GPR) have conventionally been used to identify layer thickness, delamination, and even measuring densities of pavements [27–29]. The usage of this technology in crack evaluation was limited; however, a study by Li et al. [30] integrated deep learning with GPR-based forward simulation using a 1 GHz air-coupled antenna and complex signal processing to improve detection of subgrade defects, such as looseness, voids, and cavities. The study simulated defect models, enhanced GPR images, and built a dataset to train YOLO v3 and SSD models. The findings showed that YOLO v3 achieved a higher mAP@0.5 (76.69%) compared to SSD (75.07%) on a combined dataset of standard samples and field samples. Field tests confirmed the practical applicability of the approach. The approach was successfully validated in a field test using a 450 MHz air-coupled antenna [30].

Unmanned Aerial Vehicles (UAVs) / Drones

The application of unmanned aerial vehicles (UAVs) in transportation has shown promising results in significantly improving road safety, traffic monitoring, and infrastructure management. Their effectiveness is driven by advanced computer vision and deep learning, which use the unique aerial perspective to enhance accident reconstruction, vehicle tracking, and the automated detection of infrastructure defects [31].

UAVs, combined with deep neural networks, can effectively locate major road distresses like potholes, rutting, and corrugations on unpaved roads [32]. In addition, UAV-captured video frames can be used to estimate the size of detected potholes [32]. Moreover, the substantial benefits of drone-based monitoring include reporting cost savings ranging from 13% to 94% and time savings from 37.5% to 90% compared to traditional methods [33]. However, widespread adoption faces regulatory (e.g., visual line-of-sight (VLOS) restrictions), technical (e.g., limited battery life), and social (e.g., privacy concerns) barriers [31].

Smartphones and Cameras

Smartphones can be used for collecting various images, enabling real-time capture of pavement conditions. While readily accessible and low-cost, their effectiveness can be limited by lighting and image quality without advanced processing. Mahdy et al. [34] developed deep network models for instance segmentation of six pavement distress types (longitudinal, transverse, block, and alligator cracks, bleeding, and potholes). The models were trained on a dataset of 1,040 images captured on Upper Egypt Road with a smartphone mounted on a vehicle bumper while driving. The study compared YOLOv5 to Mask R-CNN, in which the former demonstrated high object detection performance (86.8% mAP@0.5), while the latter achieved slightly better segmentation prediction. The findings suggest that YOLOv5 is more effective for object detection tasks with smaller datasets, whereas Mask R-CNN is better suited for image segmentation [34]. However, the dataset suffered from a class imbalance with a notably low number of pothole images, likely causing the models to underfit for that class. This is a typical manifestation of such an imbalance, leading to the observed performance where precision exceeded recall for potholes [34].

Another approach by Chen et al. [35] utilized smartphone accelerometer data, converting 1D vibration signals into 2D time-frequency-energy spectrograms using Wavelet Transform (WT) and Short-Time Fourier Transform (STFT) for classification with a VGG16 CNN. The WT-CNN model outperformed STFT-CNN, achieving 97.2% accuracy in distinguishing transverse cracks from manholes and normal road surfaces [35].

Object Detection Models

Popular one-stage object detection models like YOLOv3 (You Only Look Once), YOLOv5, and the latest version, YOLOv8, are widely recognized for their speed and real-time capabilities. Recent advancements have focused on enhancing their accuracy and robustness through architectural improvements. For instance, Shao et al. [36] demonstrated the YOLOv5 model's potential for crack detection and size estimation. The model was trained on images from PTZ cameras on China's Jingshi Highway. The model successfully detected and classified cracks with an 83% mAP@0.5 and estimated their size by converting pixel dimensions in low-distortion image regions into real-world measurements. Despite these promising results, the approach has notable drawbacks: its crack size estimation was validated against a very small sample of only 12 real measurements, and the method is best suited for preventive maintenance on already well-maintained roads.

Addressing common limitations like poor small-sized road defects, Sun et al. [37] integrated several key improvements to the more advanced model YOLOv8 model, such as the SPD-Conv layer and a more efficient neck structure. Their research used a subset composed of 2000 Chinese road damage images (collected using a camera mounted on a motorcycle and a

drone) from the RDD2022 dataset annotated with five types of detection objects: vertical cracks, horizontal cracks, alligator cracks, potholes, and repair. The improved YOLOv8 model achieved a 2.8% increase in mAP@0.5 (from 74.6% to 77.4%) and a 22.6% reduction in parameters compared to the baseline, while maintaining a real-time detection speed of 43 FPS. However, the model's performance under poor lighting or extreme weather conditions remains unverified. Additionally, false positives are still an issue, and the model is yet to support real-time monitoring or continuous learning for adapting to new defect types or environmental changes.

Beyond model architecture, the quality of the training data itself is a decisive factor. Fan et al. [38] conducted a systematic study comparing image preprocessing techniques like grayscaling and binarization on the performance of the YOLOv8 algorithm on various imbalanced datasets. The distresses were longitudinal cracks, transverse cracks, alligator cracks, oblique cracks, and potholes. Their primary finding was that dataset balancing had a far greater impact on performance than any preprocessing method. In fact, imbalanced datasets caused models to be biased toward dominant crack classes, whereas a custom balanced dataset of 400 images ensured more stable and generalized detection across all distress types, highlighting that small, skewed datasets lead to poor performance.

The pursuit of robust, generalizable models has led to the development of sophisticated hybrid architectures. Xie et al. [39] combined YOLOv8 with a context anchor attention mechanism (Star-CAA) that enhances feature extraction and a task-aligned dynamic (TAD) detection head that aligns classification and localization tasks dynamically. Trained on the large, multinational RDD2022 dataset (47,420 images), the model achieved a 59.7% F1-score (2.3% higher than YOLOv8-N) and a 59.9% mAP@0.5 (3.2% boost over baseline YOLOv8-N), while also being lighter and faster. Additional validation on the UAV-PDD2023 dataset confirmed strong generalization capabilities, with notable performance gains in diverse and challenging environments captured via drone imagery.

Parallel to the development of sophisticated hybrid models, significant research efforts have been dedicated to optimizing fundamental detection and segmentation approaches. For instance, a study by Xu et al. [40] directly evaluated Faster R-CNN and Mask R-CNN for crack detection without including a separate classification of crack types. The authors used data augmentation and pre-trained weights from ImageNet to train the models on a small dataset. Their results showed that while both models outperformed YOLOv3 and generalized well when tested on another dataset, Faster R-CNN provided superior bounding boxes, whereas Mask R-CNN excelled at instance segmentation. A key limitation was that both struggled with complex cracks in bifurcated, curved, or low-light conditions.

Segmentation Models

To move beyond detecting a crack to quantification of crack geometry, specialized semantic segmentation models are essential; they trade the rough bounding box for a detailed, pixel-level outline of the distress. Zhao et al. [41] introduced HWCNet, a hybrid CNN-Transformer model that excelled at preserving crack continuity in low-contrast scenarios. The model featured a specialized backbone with MobileOne Blocks for efficient local feature extraction, GAMC Blocks with dilated convolution for multiscale context capture, and SPDConv for detail-preserving downsampling. In addition, the model used a focal dice loss function that emphasizes hard-to-segment regions to deal with class imbalance between crack pixels and background. The hybrid CNN-Transformer model outperformed classical models like U-Net by achieving mAP@0.5 ranging from 82.37% to 86.95% across four datasets (Asphalt3h, GAPs, Asphalt3k, and High Way) with varying characteristics, despite remaining computationally intensive and requiring high-quality annotations.

Similarly, Yadav et al. [42] developed HCTNet for pixel-level road crack segmentation, which uses a Vision Transformer (ViT) encoder to capture long-range dependencies and a CNN decoder with a gated-attention mechanism. Evaluated on Crack500, DeepCrack, and a custom Indian dataset (1532 images), HCTNet achieved exceptional mAP@0.5 scores, with the lowest still being 97.52% on the custom dataset, demonstrating high effectiveness against noise and varied crack patterns. The ablation study showed that the gated-attention mechanism and hybrid loss function were found to significantly improve segmentation accuracy, especially against noise and varying crack patterns. This model is highly effective for automated crack segmentation, reducing reliance on manual inspection.

Whereas most models in literature excel at only one type of detection and struggle with complex field conditions like shadows and varied distress characteristics, Zhong et al. [43] proposed a solution to simultaneously achieve region-level (classification/localization) and pixel-level (segmentation) pavement distress detection through developing an end-to-end multitask fusion for pavement distress detection (PDDNet). Using a ResNet-101 backbone with multiscale feature fusion and a combined loss function including Dice loss, PDDNet was trained on a UAV-collected dataset (UAPD) of 3,151 images labeled with seven distress types. It achieved an outstanding performance with a mAP@0.5 of 81% for region-level detection and 79.5% for pixel-level segmentation, outperforming models like Mask R-CNN and U-Net and showing robustness to environmental challenges. A remaining limitation was the infrequent misclassification of large potholes as patches and difficulty distinguishing composite alligator cracks.

Table 1 Comparative summary of deep learning models for pavement distress detection, segmentation, and evaluation

Article	Architecture	Dataset	Output & Metrics	Advantages	Limitations
Xu et al. (2022) [40]	Faster R-CNN / Mask R-CNN with RoIAlign	148 smartphone images with various crack types	Bounding boxes & masksHigh Accuracy vs. YOLOv3	Works well with very few images (∼130); Outperforms YOLOv3	Small dataset; Struggles detecting complex cracks, Sensitive to settings
Zhao et al. (2025) [41]	HWCNet (CNN-Transformer hybrid with ASFF, EMA)	Asphalt3h, GAPs, Asphalt3k, High Way (varying scales & contrasts)	Pixel-level segmentationmAP@0.5: 71.99%–91.39%F1-Score: 76.83–88.78%	Strong generalization; excels on fine, low-contrast cracks	Computationally expensive; requires high-quality data
Yadav et al. (2024) [42]	HCTNet (CNN+ViT for global context + Gated-attention decoder for local details)	Crack500, DeepCrack, & a new 1532-image set	Pixel-level segmentationmAP@0.5: 97.52%IoU: 98.76%, F1: 98.69%	SOTA performance; Combines global and local features	High computational cost; Hard for real-time use
Zhong et al. (2024) [43]	PDDNet (ResNet backbone) Multiscale fusion + U-Net/YOLO-like decoder	UAPD: 3,151 UAV images, 7 distress types	Bboxes + Masks mAP@0.5: 81%-79.5% (+10.7% over Mask R-CNN)	Outperformed Mask R-CNN & U-Net significantly	Class confusion (pothole vs. patch); Struggles with shadows/rare crack types
Wen et al. (2022) [44]	PCDNet (CNN classifies patches + automatic threshold seed algorithm to refine pixels)	10,000 3D images (1 mm resolution)	Pixel-level crack mapmAP@0.5: 88.5%	Fine crack detection; Resilience to issues like shadows and surface stains.	Misidentifies road markings and sealed edges as cracks.
Abohamer et al. (2021) [45]	AlexNet CNN with automatic features learning from 3D images	850 3D images; IRI: 0.79-5.40 m/km	IRI value regressionR^2: 0.985, RMSE: 0.059	No manual features needed; Superior to ANN models	Struggles with high IRI; Small and local dataset

The development of robust crack detection and segmentation models often involves a trade-off between performance and the practicality of data preparation. While Zhong et al. pursued a high-performance, multitask solution, Wen et al. [44] prioritized data efficiency with PCD-Net, a three-phase deep learning framework model designed to deliver pixel-level crack segmentation using only patch-level annotations. This patch-level training approach significantly reduces the labor and time required for dataset elevation. The model classifies patches, identifies crack seeds, and ensures continuity with a region-growing layer. By integrating 3D elevation data, it can perform well even in the presence of shadows and stains. The model achieves a mAP@0.5 of 88.5%, a recall of 90.2%, and an F1-score of 89.3% on a diverse test set. The model excels as well at detecting fine cracks and avoids false positives from non-crack features with varying elevations. However, PCDNet struggles with exhibiting false positive detection of non-crack features with crack-like elevation profiles, such as road markings and

sealed crack edges. In addition, this model should be tested on another dataset for generalization to different pavement types and complex conditions.

Applications of deep learning extend beyond distress detection and segmentation to predicting overall pavement condition indicators. Abohamer et al. [45] employed a pre-trained AlexNet CNN to classify pavement sections into four IRI categories (poor, fair, good, and very good) using 850 3D images. For comparison, they also developed artificial neural networks (ANN) and multinomial logit (MNL) models using distress indices. The CNN significantly outperformed these traditional methods, achieving 85% accuracy (R^2 = 0.985, RMSE = 0.059) compared to approximately 58% for both ANN and MNL, confirming that distress-based predictors alone are insufficient for reliable IRI classification. Unlike preceding studies focused on detailed defect detection, Abohamer et al. [45] demonstrates CNNs' utility for high-level condition assessment. However, a key limitation emerged: the CNN's performance decreased for sections with IRI values exceeding 3.00 m/km, as high roughness variations increased prediction scatter. Notably, while the ANN and MNL models were less accurate overall, their performance remained stable across roughness levels. This contrast underscores a common challenge in deep learning applications to pavement assessment: maintaining accuracy under high-variability conditions of severely deteriorated roads.

Collectively, these studies illustrate the fundamental trade-offs between model complexity, data requirements, and task specialization in automated crack detection and segmentation. Table 1 summarizes the findings of deep learning models recently developed in literature.

Conclusion

Computer vision has firmly established itself as a cornerstone of modern pavement condition evaluation, offering transformative benefits over labor-intensive and subjective manual inspections. With innovative data acquisition methods, these systems facilitate rapid, large-scale, and non-destructive pavement assessment. The proliferation of sophisticated deep learning architectures, particularly one-stage detectors like YOLOv8 and its enhanced variants, alongside multi-task and segmentation networks, improved accuracy in detecting, classifying, and quantifying pavement distresses at both region and pixel levels.

The investment in computer vision technologies confers transformative benefits:

- *Improved Efficiency and Cost Reduction:* Automated systems drastically accelerate the inspection process, covering large road networks in a fraction of the time required for manual surveys, leading to substantial savings in labor and operational costs. Studies on UAVs report time savings of 37.5% to 90% [33].

  Objective and Quantifiable Assessment: Computer vision algorithms provide consistent, data-driven metrics for distress detection and measurement, eliminating the subjectivity and variability inherent in human visual inspection.

- *Data-Driven Maintenance Planning:* The precise, geo-referenced data generated enables predictive maintenance strategies, allowing agencies to prioritize repairs, optimize resource allocation, and extend pavement service life based on accurate condition maps.

- *Comprehensive and High-Resolution Data:* UAVs and mobile scanning systems facilitate frequent, network-level data collection, creating rich historical datasets for trend analysis and performance modeling.

Despite the remarkable progress, several critical challenges impede widespread and robust deployment:

- *Generalization and Data Limitations:* Many models struggle to generalize across different pavement types, geographic regions, or uncommon distress forms. This is often exacerbated by imbalanced training datasets, which lead to bias towards dominant classes.

- *Computational and Deployment Constraints:* The high computational cost of state-of-the-art models, particularly large CNNs and Transformers, creates a trade-off between accuracy and the feasibility of real-time processing on affordable, embedded hardware for in-vehicle systems.

- *High Cost of Data Annotation:* The supervised learning approach that dominates the field requires vast volumes of accurately labeled data. Pixel-level annotation for segmentation tasks is exceptionally time-consuming and expensive, creating a major bottleneck for model development.

- *Maintenance and Rehabilitation (M&R) Programming:* Applying AI to M&R programming is still a major research gap, requiring more effort in developing algorithms capable of solving constrained optimization problems for optimal treatment strategies.

The path forward involves creating more robust, lightweight, and environmentally resilient models. Key to this evolution will be the strategic fusion of multi-modal data, such as 2D RGB with 3D depth or GPR, and the curation of larger, more diverse datasets. Success in these areas will enable computer vision to become the central nervous system of fully autonomous, next-generation pavement management infrastructure.

References

1. Y. Xu, Z. Zhang, Review of Applications of Artificial Intelligence Algorithms in Pavement Management, Journal of Transportation Engineering, Part B: Pavements 148 (2022). https://doi.org/10.1061/jpeodx.0000373.
2. H.S. Munawar, A.W.A. Hammad, A. Haddad, C.A.P. Soares, S.T. Waller, Image-based crack detection methods: A review, Infrastructures (Basel) 6 (2021). https://doi.org/10.3390/infrastructures6080115.
3. W. Song, G. Jia, H. Zhu, D. Jia, L. Gao, Automated pavement crack damage detection using deep multiscale convolutional features, J Adv Transp 2020 (2020). https://doi.org/10.1155/2020/6412562.
4. G. Tang, C. Yin, X. Zhang, X. Liu, S. Li, Crack-Detection Method for Asphalt Pavement Based on the Improved YOLOv5, Journal of Performance of Constructed Facilities 38 (2024). https://doi.org/10.1061/jpcfev.cfeng-4615.
5. L. Ali, F.S.K. Alnajjar, W. Khan, M.A. Serhani, H. Al Jassmi, Bibliometric Analysis and Review of Deep Learning-Based Crack Detection Literature Published between 2010 and 2022, Buildings (2022). https://api.semanticscholar.org/CorpusID:247934892.
6. Y. Wang, Q. Qi, L. Sun, W. Xian, T. Ma, C. Lu, J. Zhang, An Intelligent Detection and Classification Model Based on Computer Vision for Pavement Cracks in Complicated Scenarios, Applied Sciences (2024). https://api.semanticscholar.org/CorpusID:268824673.
7. W. Li, J. Huyan, S.L. Tighe, Pavement Cracking Detection Based on Three-Dimensional Data Using Improved Active Contour Model, Journal of Transportation Engineering, Part B: Pavements 144 (2018). https://doi.org/10.1061/jpeodx.0000028.
8. X. Zhang, W. Yang, X. Tang, J. Liu, A fast learning method for accurate and robust lane detection using two-stage feature extraction with YOLO v3, Sensors (Switzerland) 18 (2018). https://doi.org/10.3390/s18124308.
9. N. Kheradmandi, V. Mehranfar, A critical review and comparative study on image segmentation-based techniques for pavement crack detection, Constr Build Mater 321 (2022). https://doi.org/10.1016/j.conbuildmat.2021.126162.
10. L. Li, B. Fang, J. Zhu, Performance Analysis of the YOLOv4 Algorithm for Pavement Damage Image Detection with Different Embedding Positions of CBAM Modules, Applied Sciences (Switzerland) 12 (2022). https://doi.org/10.3390/app121910180.
11. P. del Río-Barral, M. Soilán, S.M. González-Collazo, P. Arias, Pavement Crack Detection and Clustering via Region-Growing Algorithm from 3D MLS Point Clouds, Remote Sens (Basel) 14 (2022). https://doi.org/10.3390/rs14225866.
12. H.B. Ibrahim, M. Salah, F. Zarzoura, M. El-Mewafi, Smart monitoring of road pavement deformations from UAV images by using machine learning, Innovative Infrastructure Solutions 9 (2024). https://doi.org/10.1007/s41062-023-01315-2.
13. N.D. Hoang, Q.L. Nguyen, Fast Local Laplacian-Based Steerable and Sobel Filters Integrated with Adaptive Boosting Classification Tree for Automatic Recognition of Asphalt Pavement Cracks, Advances in Civil Engineering 2018 (2018). https://doi.org/10.1155/2018/5989246.
14. H. Zhang, Z. Qian, Y. Tan, Y. Xie, M. Li, Investigation of pavement crack detection based on deep learning method using weakly supervised instance segmentation framework, Constr Build Mater 358 (2022). https://doi.org/10.1016/j.conbuildmat.2022.129117.
15. M. Salman, S. Mathavan, K. Kamal, M. Rahman, Pavement crack detection using the Gabor filter, in: IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, 2013. https://doi.org/10.1109/ITSC.2013.6728529.
16. H. Ahmad, M. Salman, S. Hussain, K. Khurshid, Automation of Optimized Gabor Filter Parameter Selection for Road Cracks Detection, International Journal of Advanced Computer Science and Applications 7 (2016). https://doi.org/10.14569/ijacsa.2016.070339.
17. J. Wang, X. Li, Y. Xu, Z. Zhou, W. Wu, Z. Li, Optimizing CNN for pavement distress detection via edge-enhanced multi-scale feature fusion, PLoS One 20 (2025) e0319299. https://doi.org/10.1371/journal.pone.0319299.
18. P. Kamranfar, D. Lattanzi, A. Shehu, S. Stoffels, Pavement Distress Recognition via Wavelet-Based Clustering of Smartphone Accelerometer Data, Journal of Computing in Civil Engineering 36 (2022). https://doi.org/10.1061/(asce)cp.1943-5487.0001022.
19. M. Quintana, J. Torres, J.M. Menéndez, A simplified computer vision system for road surface inspection and maintenance, IEEE Transactions on Intelligent Transportation Systems 17 (2016). https://doi.org/10.1109/TITS.2015.2482222.
20. S. Mathavan, K. Kamal, M. Rahman, A Review of Three-Dimensional Imaging Technologies for Pavement Distress Detection and Measurements, IEEE Transactions on Intelligent Transportation Systems 16 (2015). https://doi.org/10.1109/TITS.2015.2428655.
21. Y.-A. Hsieh, Y.J. Tsai, Machine Learning for Crack Detection: Review and Model Performance Comparison, Journal of Computing in Civil Engineering 34 (2020). https://doi.org/10.1061/(asce)cp.1943-5487.0000918.
22. E. Kaartinen, K. Dunphy, A. Sadhu, LiDAR-Based Structural Health Monitoring: Applications in Civil Infrastructure Systems, Sensors 22 (2022). https://doi.org/10.3390/s22124610.
23. T. Yi, D. Ting, Z. Jingyu, Z. Zhixiang, LiDAR-Based Automatic Pavement Distress Detection and Management Using Deep Learning and BIM, J Constr Eng Manag 150 (2024) 04024069. https://doi.org/10.1061/JCEMD4.COENG-14358.
24. P. Asadi, H. Mehrabi, A. Asadi, M. Ahmadi, Deep convolutional neural networks for pavement crack detection using an inexpensive global shutter rgb-d sensor and arm-based single-board computer, Transp Res Rec 2675 (2021). https://doi.org/10.1177/03611981211004974.
25. Y.-T. Huang, M.R. Jahanshahi, F. Shen, T.G. Mondal, Deep Learning–Based Autonomous Road Condition Assessment Leveraging Inexpensive RGB and Depth Sensors and Heterogeneous Data Fusion: Pothole Detection and Quantification, Journal of Transportation Engineering, Part B: Pavements 149 (2023). https://doi.org/10.1061/jpeodx.pveng-1194.
26. A. Mahmoudzadeh, A. Golroo, M.R. Jahanshahi, S.F. Yeganeh, Estimating pavement roughness by fusing color and depth data obtained from an inexpensive RGB-D sensor, Sensors (Switzerland) 19 (2019). https://doi.org/10.3390/s19071655.
27. S. Wang, S. Zhao, I.L. Al-Qadi, Continuous real-time monitoring of flexible pavement layer density and thickness using ground penetrating radar, NDT and E International 100 (2018). https://doi.org/10.1016/j.ndteint.2018.08.005.

28. S. Wang, Z. Leng, Z. Zhang, X. Sui, Automatic asphalt layer interface detection and thickness determination from ground-penetrating radar data, Constr Build Mater 357 (2022) 129434. https://doi.org/10.1016/J.CONBUILDMAT.2022.129434.
29. W. He, W. Wai-Lok Lai, X. Sui, A. Giannopoulos, Delamination characterization in thin asphalt pavement structure using dispersive GPR data, Constr Build Mater 402 (2023) 132834. https://doi.org/10.1016/J.CONBUILDMAT.2023.132834.
30. P. Li, M. Bai, X. Li, C. Liu, Research on the Forward Simulation and Intelligent Detection of Defects in Highways Using Ground-Penetrating Radar, Applied Sciences 14 (2024). https://doi.org/10.3390/app142210183.
31. F. Outay, H.A. Mengash, M. Adnan, Applications of unmanned aerial vehicle (UAV) in road safety, traffic and highway infrastructure management: Recent advances and challenges, Transp Res Part A Policy Pract 141 (2020). https://doi.org/10.1016/j.tra.2020.09.018.
32. L. Loures, E. Azar, Condition Assessment of Unpaved Roads Using Low-Cost Computer Vision–Based Solutions, Journal of Transportation Engineering, Part B: Pavements 149 (2023). https://doi.org/10.1061/jpeodx.pveng-1006.
33. A. Taraneh, B. Raj, T.D. D, Drones for Road Condition Monitoring: Applications and Benefits, Journal of Transportation Engineering, Part B: Pavements 151 (2025) 04024055. https://doi.org/10.1061/JPEODX.PVENG-1559.
34. K. Mahdy, A. Zekry, M. Moussa, A. Mohamed, H. Mahdy, M. Elhabiby, Pavement distress instance segmentation using deep neural networks and low-cost sensors, Innovative Infrastructure Solutions 9 (2024). https://doi.org/10.1007/s41062-023-01308-1.
35. C. Chen, H. Seo, Y. Zhao, A novel pavement transverse cracks detection model using WT-CNN and STFT-CNN for smartphone data analysis, International Journal of Pavement Engineering 23 (2022). https://doi.org/10.1080/10298436.2021.1945056.
36. C. Shao, X. Feng, J. Li, Crack Detection and Measurement Using PTZ Camera–Based Image Processing Method on Expressways, Journal of Transportation Engineering, Part B: Pavements 148 (2022). https://doi.org/10.1061/jpeodx.0000390.
37. Z. Sun, L. Zhu, S. Qin, Y. Yu, R. Ju, Q. Li, Road Surface Defect Detection Algorithm Based on YOLOv8, Electronics (Basel) 13 (2024). https://doi.org/10.3390/electronics13122413.
38. L. Fan, S. Tang, M.K.A. b. M. Ariffin, M.I.S. Ismail, X. Wang, Impact of Image Preprocessing and Crack Type Distribution on YOLOv8-Based Road Crack Detection, Sensors 25 (2025). https://doi.org/10.3390/s25072180.
39. Y. Xie, D. Du, Z. Wang, Y. Liu, M. Bi, YOLO-SFT: Road Damage Detection Algorithm Based on Feature Diffusion, Journal of Transportation Engineering, Part B: Pavements 151 (2025) 4025004. https://doi.org/10.1061/JPEODX.PVENG-1726.
40. X. Xu, M. Zhao, P. Shi, R. Ren, X. He, X. Wei, H. Yang, Crack Detection and Comparison Study Based on Faster R-CNN and Mask R-CNN, Sensors 22 (2022). https://doi.org/10.3390/s22031215.
41. Q. Zhao, Z. Liu, F. Meng, Z. Ma, A. Tuerhong, G. Guo, HWCNet: A Semantic Segmentation Model for Highway Asphalt Pavement Cracks, Journal of Transportation Engineering, Part B: Pavements 151 (2025) 4025035. https://doi.org/10.1061/JPEODX.PVENG-1635.
42. D.P. Yadav, B. Sharma, S. Chauhan, F. Amin, R. Abbasi, Enhancing Road Crack Localization for Sustainable Road Safety Using HCTNet, Sustainability 16 (2024). https://doi.org/10.3390/su16114409.
43. J. Zhong, M. Zhang, Y. Ma, R. Xiao, G. Cheng, B. Huang, A Multitask Fusion Network for Region-Level and Pixel-Level Pavement Distress Detection, Journal of Transportation Engineering, Part B: Pavements 150 (2024) 4024002. https://doi.org/10.1061/JPEODX.PVENG-1433.
44. T. Wen, H. Lang, S. Ding, J.J. Lu, Y. Xing, PCDNet: Seed Operation–Based Deep Learning Model for Pavement Crack Detection on 3D Asphalt Surface, Journal of Transportation Engineering, Part B: Pavements 148 (2022) 4022023. https://doi.org/10.1061/JPEODX.0000367.
45. H. Abohamer, M. Elseifi, N. Dhakal, Z. Zhang, C.N. Fillastre, Development of a Deep Convolutional Neural Network for the Prediction of Pavement Roughness from 3D Images, Journal of Transportation Engineering, Part B: Pavements 147 (2021) 4021048. https://doi.org/10.1061/JPEODX.0000310.

Chapter 12
Vortex Generation by Phase-Controlled Structure-Borne Traveling Waves in a Submerged Beam: A PIV/PTV Fluid–Structure Interaction Study

Rachel Cole, Samikhshak Gupta, Rohit S. Pandhare, and Vijaya V. N. Sriram Malladi

Abstract This work investigates the transfer of mechanical energy from structure-borne traveling waves (TWs) in a slender beam to the surrounding fluid, with a focus on the resulting vortex dynamics. An Aluminum 6061 beam was immersed in water and excited using dual electrodynamic shakers with a controlled phase offset to generate steady traveling waves. Modal testing was employed to identify natural frequencies and establish traveling-wave conditions, while particle tracking velocimetry (PTV) captured near-body flow fields. The PTV measurements resolved coherent vortex formation and fluid motion around the beam, revealing alternating vortices that propagated along its length and induced linear flow within adjacent channels. Analysis showed a consistent relationship between wave kinematics and fluid response, with approximately four vortices forming per excitation wavelength. These results establish an experimental framework linking traveling-wave parameters to circulation and fluid transport. The findings highlight the potential of programmable structural waves as a low-maintenance mechanism for fluid mixing, cooling, and propulsion, offering an alternative to mechanically complex pumps, fans, or rotors.

Keywords traveling wave · PTV · SLDV · vortex formation · modal analysis

Introduction

The transportation sector is undergoing a major transition as governments enforce stricter emission standards and consumers demand cleaner alternatives to fossil-fuel vehicles. Electric vehicles (EVs), including hybrid (HEVs) and battery electric vehicles (BEVs), are central to this shift. Since batteries account for nearly one-third of EV costs, ongoing reductions in price are a key enabler of adoption. Forecasts predict EVs will comprise up to 32% of global vehicle sales by 2030, with BEVs dominating the share. This growth intensifies demand for advanced batteries that offer high energy density, fast charging, low weight, and long cycle life. Lithium-ion batteries (LIBs), the prevailing technology, satisfy many of these requirements but remain highly vulnerable to temperature variations [1–4]. Heat generation during charge–discharge cycles [5–7] and exposure to extreme climates accelerate degradation and raise safety risks, most notably thermal runaway [8–10]. Consequently, robust battery thermal management systems (BTMS) are essential to ensure the safety, durability, and efficiency of EVs. Current BTMS strategies are typically classified as air cooling [11–14], single-phase liquid cooling [15–20], or two-phase cooling [21–23]. Early EVs such as the Toyota Prius and Honda Insight employed air cooling, while later designs such as the Chevrolet Volt and Tesla Model S adopted liquid cooling to improve heat dissipation [20–28]. More recently, immersion cooling with dielectric fluids (e.g., mineral oils) has demonstrated superior thermal performance [29–31], lowering peak cell temperatures and stabilizing fluctuations [32]. However, liquid-based systems require auxiliary power to drive pumps, blowers, and fans. These parasitic loads reduce both vehicle range and efficiency. The challenge is therefore not only to regulate temperature effectively but also to do so with minimal additional energy demand.

Rachel Cole · Samikhshak Gupta · Vijaya V. N. Sriram Malladi
Vibrations, Intelligent Testing, & Active Learning of Structures (ViTALS) Group, Department of Mechanical and Aerospace Engineering, e-mail: recole@mtu.edu; sgupta14@mtu.edu; smalladi@mtu.edu

Rohit S. Pandhare
Complex Fluids and Active Matter (CFAM) Group, Department of Mechanical and Aerospace Engineering, Michigan Technological University, Houghton, MI 49931

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

To overcome this limitation, researchers have explored energy recovery strategies that repurpose wasted vibrational and thermal energy within vehicles. Energy harvesting approaches—including piezoelectric absorbers [33–40], thermoelectric generators [41–45], and shape memory alloy engines [46–48]–seek to convert otherwise wasted energy into usable power. While current systems face limitations in efficiency and scalability, they suggest a promising direction: integrating harvested energy with fluid circulation to reduce dependence on electrically powered pumps. Such strategies could enhance BTMS efficiency, extend LIB service life, and improve overall EV performance. Building on this concept, the present work introduces a mechanism that harnesses vibrational noise for fluid circulation, thereby reducing parasitic energy losses and advancing sustainable electric mobility [49–54].

Conventional marine propulsors face similar efficiency and maintenance challenges, as they depend on rotating components. A bio-inspired alternative is thrust generation via undulatory motion of finite structures, in which structure-borne traveling waves (TWs) shed momentum into the fluid to produce net propulsion [55–57]. In elastic media, wave behavior depends strongly on boundary conditions: traveling waves transport energy, whereas standing waves localize it spatially. Single-point excitation in finite structures typically excites reflections and standing modes [58, 59]. By contrast, multiple inputs with controlled phase offsets can suppress reflections and bias energy transport, producing hybrid or near-pure TW fields. Although fully reflection-free TWs remain challenging in bounded domains [60], recent advances have demonstrated controlled TW synthesis across various geometries, including strings [58, 61], beams [62–65], cylinders [59], and plates [66–69]. For example, Malladi et al. developed and experimentally validated a model for TW synthesis in one-dimensional beams under varied boundary conditions using piezoelectric actuation [65], while more recent work has demonstrated distributed actuation in slender plates [70]. The ability to tailor TW kinematics in solids has clear implications for underwater propulsion [71, 68, 55] and for emerging applications such as thermal management, offering a pathway to controllable thrust and circulation without reliance on moving parts.

The central gap lies in establishing a measurement-based understanding of how imposed TW motion couples with the surrounding fluid to generate and organize vortices that mediate momentum exchange. In this study, steady, structure-borne TWs are excited in a one-dimensional beam, and the resulting fluid–structure interactions are characterized using particle tracking velocimetry (PTV). These measurements resolve near-body velocity fields and vortex dynamics, enabling direct correlations between TW parameters (frequency, amplitude, wavenumber, and phase speed) and vortex formation at the beam surface. The objective is to quantify circulation, vortex motion, and flow speeds to determine how TW motion governs fluid velocity and propulsive performance. Ultimately, these insights support the development of low-maintenance cooling and propulsion systems that harness programmable structural waves in place of mechanically complex pumps, fans, or rotors.

Underwater Modal Analysis

The experimental setup for modal analysis of a one-dimensional aluminum beam is shown in Figure 1. The test specimen was an Aluminum 6061 beam with dimensions of $381\text{mm} \times 25.4\text{mm} \times 0.794\text{mm}$. The beam was suspended at both ends using 3D-printed clips bonded with epoxy. Each clip featured a narrow attachment section to minimize fluid disturbance and widened at the top to allow bolting to metal support bars. The support bars were connected to horizontally mounted Modal Shop Smart Shakers to provide excitation. The experiments were conducted in a water tank with internal dimensions of $444.5\text{mm} \times 444.5\text{mm} \times 76.2\text{mm}$, filled to a depth of 114.3mm. The beam was positioned laterally at the tank center, with its lower edge 3mm above the tank floor. Proper alignment between the shakers and the beam was maintained to avoid introducing residual stresses during testing [72].

The supporting frame was constructed from 80:20 aluminum extrusions to ensure rigidity and stability of the test apparatus. Response measurements were obtained using a Polytec OFV-353 laser vibrometer sensor head coupled with an OFV-040 scan unit. The scan unit deflected the laser beam within an angular range of $\pm 15^\circ$ in both x and y axes, controlled through analog voltages (0 to $\pm 5\text{V}$) supplied by an OFV-042 scan controller. Data acquisition was managed by an OFV-3001 vibrometer controller, with velocity measurements taken at 75 uniformly distributed points along the beam. To enhance signal quality, reflective tape was applied along the beam surface. Excitation was provided by the suspended shakers using a chirp signal ranging from 0.1–300Hz with an amplitude of 0.0375V Figure 1. The vibrometer acquired data with 10 averages at each measurement location. Sampling was performed at 17067Hz over a duration of 5s, with a Hanning window applied for spectral analysis. A two-force input strategy was implemented, followed by a multi-input multi-output (MIMO) force-feedback test using the same chirp excitation. Data acquisition employed an NI cDAQ-9174 chassis with NI 9263 and NI 9234 I/O modules. The frequency response functions (FRFs) were computed between the shaker input voltages and the corresponding vibrometer velocity outputs and are shown in Figure 2.

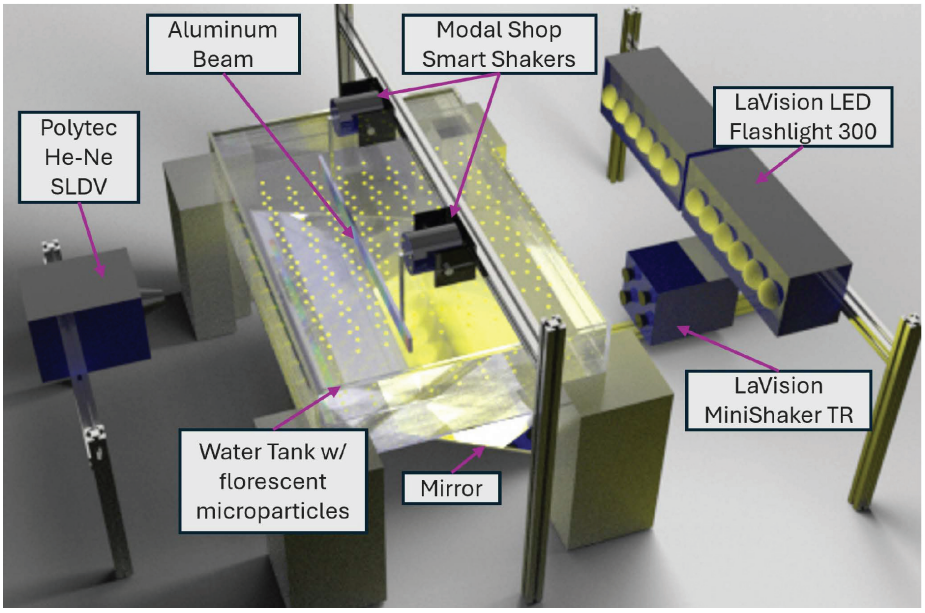
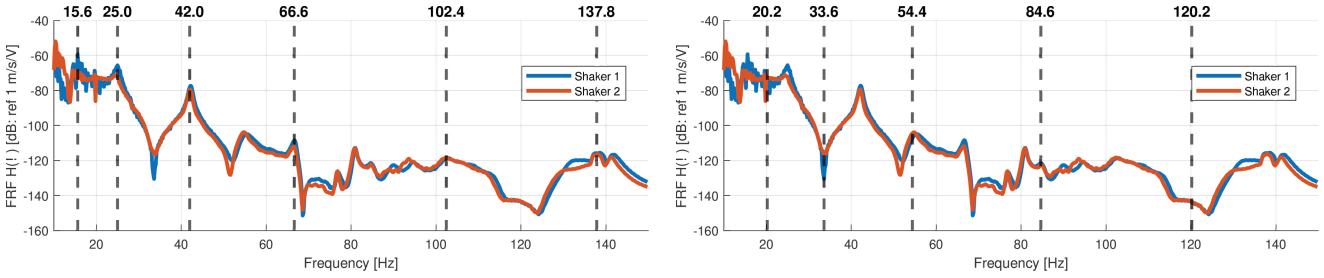


Fig. 1 Schematic Diagram for experimental setup for modal analysis and particle tracking velocimetry of the submerged Al-Beam



(a) Measured Frequency Response Functions (FRFs) of the aluminum beam with natural frequencies highlighted

(b) Measured Frequency Response Functions (FRFs) of the aluminum beam with traveling wave frequencies highlighted

Fig. 2 Measured Frequency Response Function (FRF) of the submerged aluminum beam

Traveling Wave Generation

Previous studies have shown that elastic structures can support traveling waves when excited by two actuators operated out of phase [73, 65, 74, 75]. The key requirement is to drive the actuators at a mean frequency between adjacent resonances while maintaining a 90° phase shift. This concept was experimentally demonstrated on an aluminum beam using two modal shakers driven at equal amplitudes and a controlled phase offset of 90°. The superposition of their responses produced a structure-borne traveling wave (SBTW). The excitation applied to the two shakers can be expressed as $V_1 = A\sin(2\pi ft)$ and $V_2 = A\sin(2\pi ft \pm \phi)$, leading to a combined input of

$$\textbf{SBTW} = A\sin(2\pi ft) \;\pm\; A\sin(2\pi ft \pm \phi), \tag{1}$$

where A is the excitation amplitude, f is frequency, t is time, and ϕ is the phase difference, ideally 90° for a pure traveling wave. The quality of the generated wave depends primarily on the imposed phase relationship between actuators. In practice, an SBTW often consists of both traveling and standing wave components; reducing the standing contribution enhances wave purity. Malladi et al. [55] proposed a cost function approach to quantify this effect by systematically assessing the strength of the traveling component. The natural frequencies are presented in Figure 2a, while the traveling wave frequencies are shown in Figure 2b and summarized in Table 1.

Table 1 Identified frequencies and corresponding modal or traveling-wave responses.

Frequency	Response Type
15.6 Hz	1st mode shape
20.3 Hz	Traveling wave
25.0 Hz	2nd mode shape
33.55 Hz	Traveling wave
42.0 Hz	3rd mode shape
54.35 Hz	Traveling wave
66.6 Hz	4th mode shape
84.55 Hz	Traveling wave
102.4 Hz	5th mode shape
120.15 Hz	Traveling wave
137.8 Hz	6th mode shape

Particle Tracking Velocimetry

To characterize the flow field generated by the vibrating beam, particle tracking velocimetry (PTV) experiments were conducted. The measurement system consisted of a LaVision MiniShaker TR/2M equipped with four stereo cameras operating at 170 Hz, synchronized with a pulsed LaVision LED-Flashlight 300 for illumination. A 45° mirror was mounted beneath

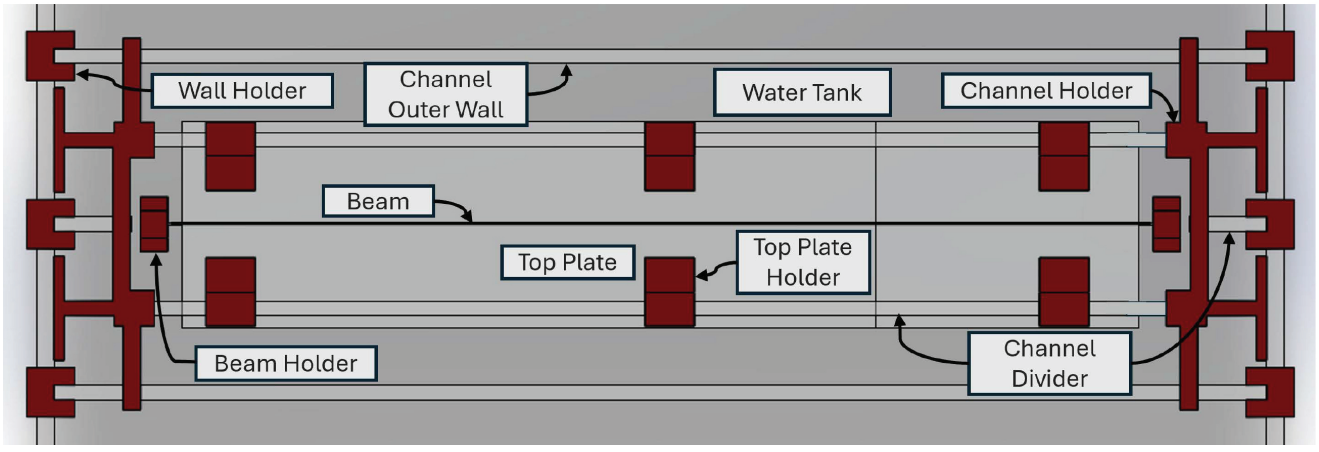


Fig. 3 Top view of underwater beam 19.05 mm channel setup.

the water tank to provide a view normal to the beam's plane of motion. Fluorescent polyethylene microspheres (90–106 μm) were introduced into the water as tracer particles Figure 1. The imaging system was calibrated using a LaVision 309-15 3-D calibration plate illuminated by seven LED light bars, while the water tank was calibrated with the LaVision DaVis volume calibration system, focusing on the water in the space 8 mm below and 9 mm above the center of the beam. To minimize background interference, all overhead lighting was turned off during image acquisition.

Experiments were performed in a double 19.05 mm water channel, with a total width of 117.6 mm from outer wall to outer wall, designed to confine the flow field generated by the vibrating beam Figure 3. The water depth was set to 45 mm, and a top plate was positioned 3 mm above the beam. Before each run, the water was stirred to evenly distribute tracer particles and then allowed to settle to eliminate external disturbances.

During testing, the beam was excited with a 0.4 V sine wave applied to both shakers, with one shaker driven 90° out of phase relative to the other. The system was allowed to stabilize for 10 s before data acquisition began. A total of 600 images were recorded over 3.52 s. Post-processing of the image sets included time-filter subtraction, image pre-processing with masking to isolate the water channel, Shake-the-Box particle tracking, and fine-scale reconstruction.

Results

The focus of this analysis was on excitation frequencies up to 150 Hz. Within this range, the modal analysis identified six natural frequencies for the beam, as shown in the FRF plot in Figure 2. Using the methodology described previously, five traveling waves were also determined, each with the potential to drive fluid motion through the channel setup (Table 1). For the present study, attention was directed to a single traveling wave at 20.3 Hz, where the flow fields around the beam were mapped using the PTV system. Upon visual inspection of both the physical system and the raw PTV images, distinct vortex motion was observed in the inner channel, while linear particle movement was primarily noted in the outer channel. These qualitative observations were later confirmed through processing with the Shake-the-Box algorithm, which detects and tracks the particles over time, as shown in Figure 4. To convert the particle tracks into vector field data, a fine-scale reconstruction was performed on the four regions of the double channel, designated as the inner and outer portions of the upper and lower channels Figure 5. Each of these regions was processed separately to avoid interpolation artifacts across the channel dividers or at the beam location.

In the fine-scale reconstruction, vortices in the Z-direction are clearly visible within the inner channels. These vortices are represented by alternating red and blue regions forming along the beam length, corresponding to positive and negative vorticity, respectively. In contrast, the outer channels display a nearly uniform green field, indicating an absence of strong vortical activity. This observation is consistent with the Shake-the-Box velocity vectors, which show primarily linear particle motion in the outer channels compared to the well-defined rotational motion in the inner channels.

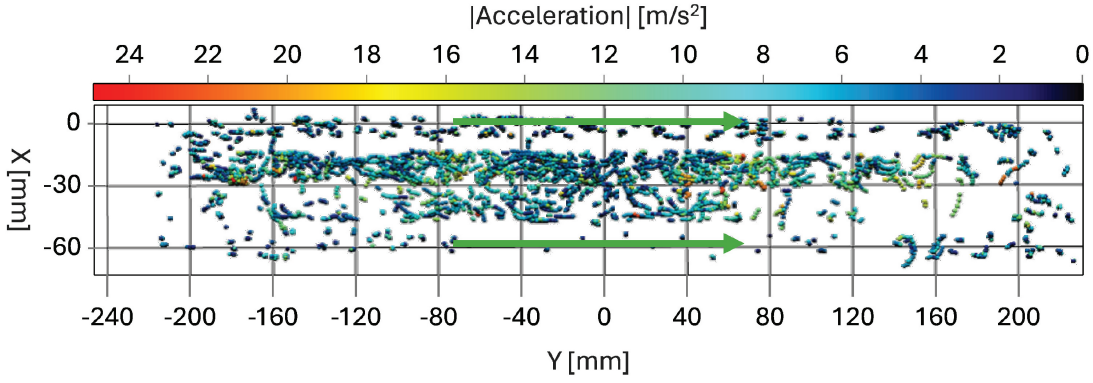
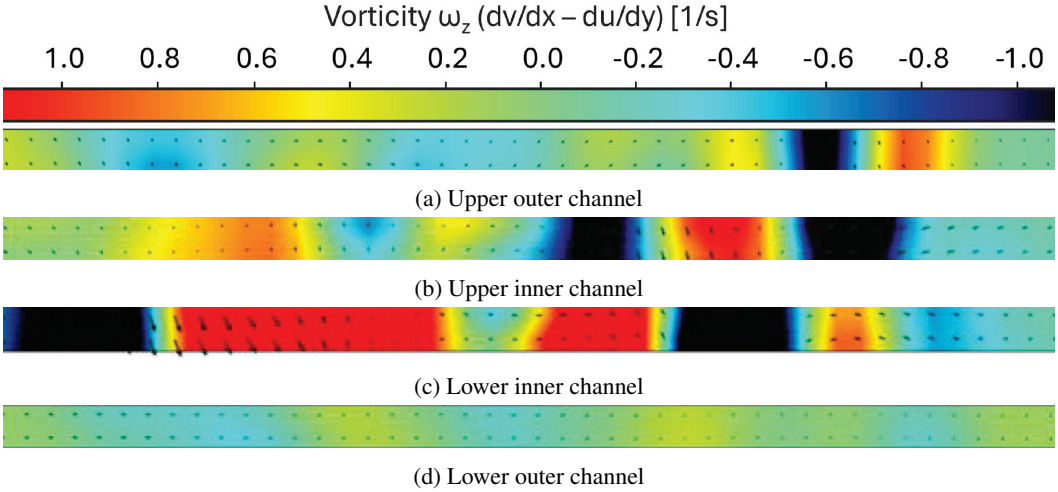


Fig. 4 Shake-the-Box reconstruction of particle motion at 20.3 Hz traveling wave in a 19.05 mm channel



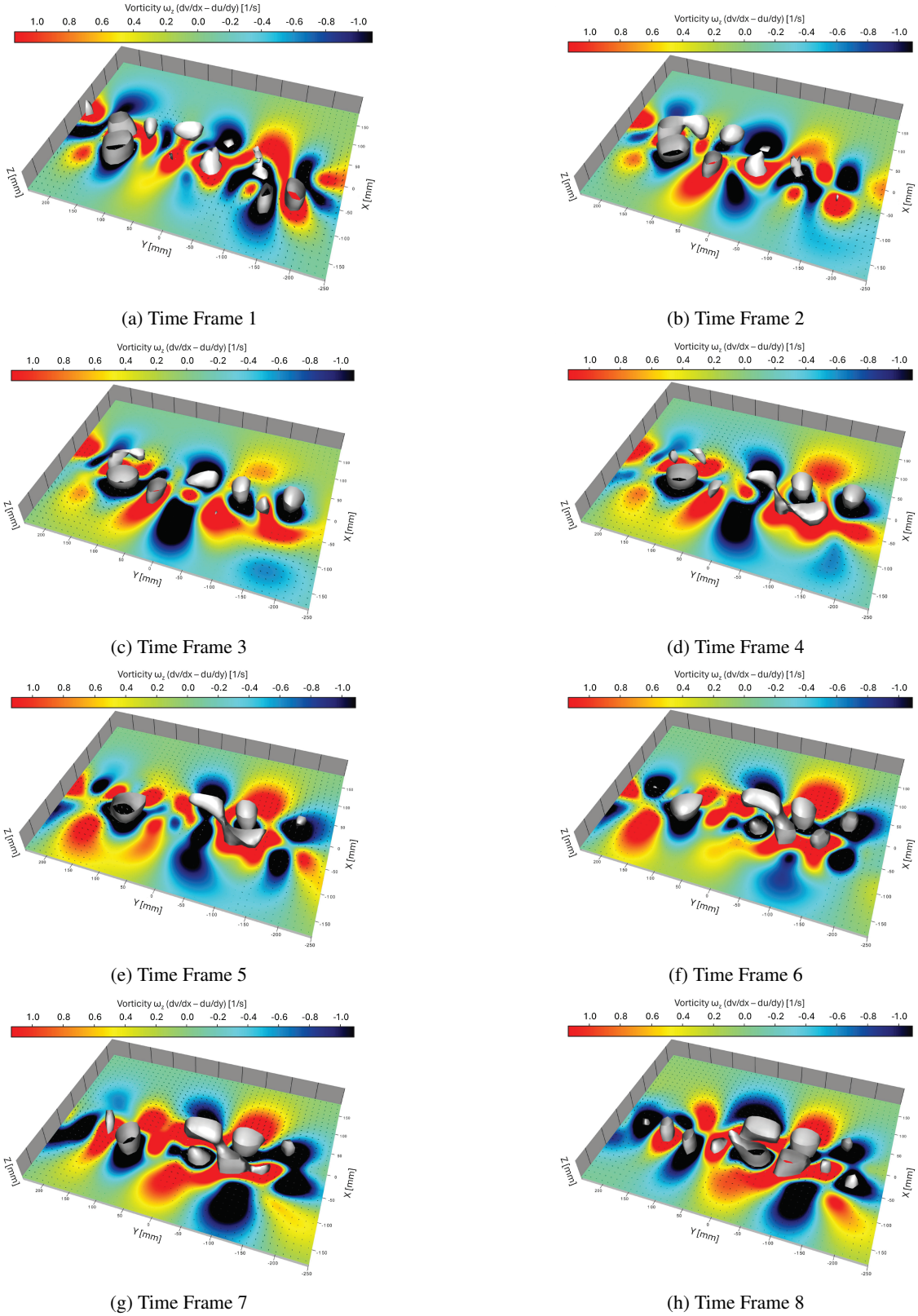
(a) Upper outer channel

(b) Upper inner channel

(c) Lower inner channel

(d) Lower outer channel

Fig. 5 Fine-scale reconstruction of the 20.3 Hz traveling wave, showing inner and outer regions of the upper and lower channels.



(a) Time Frame 1

(b) Time Frame 2

(c) Time Frame 3

(d) Time Frame 4

(e) Time Frame 5

(f) Time Frame 6

(g) Time Frame 7

(h) Time Frame 8

Fig. 6 Fine-scale reconstruction of the 20.3 Hz traveling wave, showing iso-surfaces across time.

The PTV results demonstrate that two-shaker excitation of a beam in water can generate traveling waves. In the channel configuration, the traveling wave excitation induced vortex structures whose linear momentum resulted in a sustained circular flow throughout the channel system.

Conclusion

Previous studies have reported fluid and particle motion induced by traveling waves, but the mechanisms underlying this motion remain poorly understood. The present experimental model provides a framework for developing a fluid–structure interaction (FSI) model to examine how structure-borne traveling waves couple with the surrounding fluid and drive flow. The fine-scale reconstruction reveals the formation of vortices around the beam. Close inspection over time shows alternating vortices propagating along the beam in the y-direction, producing linear motion in the outer channels. A sequence of eight frames from the dataset is presented in Figure 6 to illustrate this dynamic evolution.

Analysis of the reconstructions further indicates a clear relationship between the traveling wave wavelength and vortex formation: approximately four vortices are generated per excitation wavelength. This observation highlights the role of traveling waves in organizing flow structures and motivates further investigation into FSI models to better characterize vortex number, propagation speed, and related phenomena.

Acknowledgments The author (RC) gratefully acknowledges support from the DoD through the DoD SMART Scholarship. The authors (VVNSM and SG) also acknowledge support from the National Science Foundation under grant CMMI-2301776 for funding the research presented in this paper.

Declaration of generative AI and AI-assisted technologies in the writing process

While preparing this work, the author(s) used ChatGPT to edit the language. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

1. Liu, Y., Zhu, Y., and Cui, Y. "Challenges and opportunities towards fast-charging battery materials". *Nature Energy*, 4(7):540–550 (2019).
2. Keyser, M., Pesaran, A., Li, Q., Santhanagopalan, S., Smith, K., Wood, E., Ahmed, S., Bloom, I., Dufek, E., Shirk, M., et al. "Enabling fast charging–battery thermal considerations". *Journal of Power Sources*, 367:228–236 (2017).
3. Xu, K. "Nonaqueous liquid electrolytes for lithium-based rechargeable batteries". *Chemical reviews*, 104(10):4303–4418 (2004).
4. Lin, D., Liu, Y., and Cui, Y. "Reviving the lithium metal anode for high-energy batteries". *Nature nanotechnology*, 12(3):194–206 (2017).
5. Olabi, A., Maghrabie, H.M., Adhari, O.H.K., Sayed, E.T., Yousef, B.A., Salameh, T., Kamil, M., and Abdelkareem, M.A. "Battery thermal management systems: Recent progress and challenges". *International Journal of Thermofluids*, 15:100171 (2022).
6. Spotnitz, R. and Franklin, J. "Abuse behavior of high-power, lithium-ion cells". *Journal of power sources*, 113(1):81–100 (2003).
7. Al Hallaj, S. and Selman, J. "A novel thermal management system for electric vehicle batteries using phase-change material". *Journal of the Electrochemical Society*, 147(9):3231 (2000).
8. Kong, L., Li, C., Jiang, J., and Pecht, M.G. "Li-ion battery fire hazards and safety strategies". *Energies*, 11(9):2191 (2018).
9. Lu, L., Han, X., Li, J., Hua, J., and Ouyang, M. "A review on the key issues for lithium-ion battery management in electric vehicles". *Journal of power sources*, 226:272–288 (2013).
10. Wang, Q., Ping, P., Zhao, X., Chu, G., Sun, J., and Chen, C. "Thermal runaway caused fire and explosion of lithium ion battery". *Journal of power sources*, 208:210–224 (2012).
11. Choi, Y.S. and Kang, D.M. "Prediction of thermal behaviors of an air-cooled lithium-ion battery system for hybrid electric vehicles". *Journal of Power Sources*, 270:273–280 (2014).
12. Bandhauer, T.M., Garimella, S., and Fuller, T.F. "A critical review of thermal issues in lithium-ion batteries". *Journal of the electrochemical society*, 158(3):R1 (2011).
13. Yang, N., Zhang, X., Li, G., and Hua, D. "Assessment of the forced air-cooling performance for cylindrical lithium-ion battery packs: A comparative analysis between aligned and staggered cell arrangements". *Applied thermal engineering*, 80:55–65 (2015).
14. Sabbah, R., Kizilel, R., Selman, J., and Al-Hallaj, S. "Active (air-cooled) vs. passive (phase change material) thermal management of high power lithium-ion packs: Limitation of temperature rise and uniformity of temperature distribution". *Journal of power sources*, 182(2):630–638 (2008).
15. Pesaran, A., Keyser, M., and Burch, S. "An approach for designing thermal management systems for electric and hybrid vehicle battery packs". Technical report, National Renewable Energy Laboratory, Golden, CO (US) (1999).
16. Deng, Y., Feng, C., Zhu, H., Chen, J., Wen, M., Yin, H., et al. "Effects of different coolants and cooling strategies on the cooling performance of the power lithium ion battery system: A review". *Applied Thermal Engineering*, 142:10–29 (2018).
17. Roe, C., Feng, X., White, G., Li, R., Wang, H., Rui, X., Li, C., Zhang, F., Null, V., Parkes, M., et al. "Immersion cooling for lithium-ion batteries–a review". *Journal of Power Sources*, 525:231094 (2022).

18. Rao, Z. and Wang, S. "A review of power battery thermal energy management". *Renewable and Sustainable Energy Reviews*, 15(9):4554–4571 (2011).
19. Bernardi, D., Pawlikowski, E., and Newman, J. "A general energy balance for battery systems". *Journal of the electrochemical society*, 132(1):5 (1985).
20. Chen, D., Jiang, J., Kim, G.H., Yang, C., and Pesaran, A. "Comparison of different cooling methods for lithium ion battery cells". *Applied thermal engineering*, 94:846–854 (2016).
21. Fan, S. and Duan, F. "A review of two-phase submerged boiling in thermal management of electronic cooling". *International Journal of Heat and Mass Transfer*, 150:119324 (2020).
22. Ataur, R., Hawlader, M.N.A., and Khalid, H. "Two-phase evaporative battery thermal management technology for evs/hevs". *International journal of automotive technology*, 18(5):875–882 (2017).
23. Shen, M. and Gao, Q. "System simulation on refrigerant-based battery thermal management technology for electric vehicles". *Energy Conversion and Management*, 203:112176 (2020).
24. Zolot, M., Pesaran, A.A., and Mihalic, M. "Thermal evaluation of toyota prius battery pack". Technical report, SAE Technical paper (2002).
25. Zolot, M.D., Kelly, K., Keyser, M., Mihalic, M., Pesaran, A., and Hieronymus, A. "Thermal evaluation of the honda insight battery pack". Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States) (2001).
26. Kelly, K.J., Mihalic, M., and Zolot, M. "Battery usage and thermal performance of the toyota prius and honda insight during chassis dynamometer testing". In *Seventeenth Annual Battery Conference on Applications and Advances. Proceedings of Conference (Cat. No. 02TH8576)*, pages 247–252. IEEE (2002).
27. Ji, Y. and Wang, C.Y. "Heating strategies for li-ion batteries operated from subzero temperatures". *Electrochimica Acta*, 107:664–674 (2013).
28. Wang, Q., Jiang, B., Li, B., and Yan, Y. "A critical review of thermal management models and solutions of lithium-ion batteries for the development of pure electric vehicles". *Renewable and sustainable energy reviews*, 64:106–128 (2016).
29. Pambudi, N.A., Sarifudin, A., Firdaus, R.A., Ulfa, D.K., Gandidi, I.M., and Romadhon, R. "The immersion cooling technology: Current and future development in energy saving". *Alexandria Engineering Journal*, 61(12):9509–9527 (2022).
30. Murshed, S.S. and De Castro, C.N. "A critical review of traditional and emerging techniques and fluids for electronics cooling". *Renewable and Sustainable Energy Reviews*, 78:821–833 (2017).
31. Khalaj, A.H. and Halgamuge, S.K. "A review on efficient thermal management of air-and liquid-cooled data centers: From chip to the cooling system". *Applied energy*, 205:1165–1188 (2017).
32. Pesaran, A. and Kim, G. "Battery thermal management system design modeling". Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States) (2006).
33. Neighborgall, C.R., Kothari, K., Sriram Malladi, V., Tarazaga, P., Paruchuri, S.T., and Kurdila, A. "Shaping the frequency response function (frf) of a multi-degree-of-freedom (mdof) structure using arrays of tuned vibration absorbers (tva)". In *Topics in Modal Analysis & Testing, Volume 8: Proceedings of the 37th IMAC, A Conference and Exposition on Structural Dynamics 2019*, pages 317–326. Springer (2019).
34. Paruchuri, S.T., Sterling, J., Malladi, V.V.S., Kurdila, A., Vignola, J., and Tarazaga, P. "Passive piezoelectric subordinate oscillator arrays". *Smart Materials and Structures*, 28(8):085046 (2019).
35. Sterling, J., Paruchuri, S.T., Tarazaga, P., Vignola, J., Kurdila, A., Malladi, V.V.S., and Ryan, T. "Piezoelectric subordinate oscillator arrays: Performance recovery and robustness to uncertainty". In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, volume 59285, page V008T10A043. American Society of Mechanical Engineers (2019).
36. Wang, Y. and Inman, D.J. "A survey of control strategies for simultaneous vibration suppression and energy harvesting via piezoceramics". *Journal of Intelligent Material Systems and Structures*, 23(18):2021–2037 (2012).
37. Sodano, H.A., Inman, D.J., and Park, G. "A review of power harvesting from vibration using piezoelectric materials". *Shock and Vibration Digest*, 36(3):197–206 (2004).
38. Junior, C.D.M., Erturk, A., and Inman, D.J. "An electromechanical finite element model for piezoelectric energy harvester plates". *Journal of Sound and Vibration*, 327(1-2):9–25 (2009).
39. Erturk, A. and Inman, D.J. *Piezoelectric energy harvesting*. John Wiley & Sons (2011).
40. Priya, S. "Advances in energy harvesting using low profile piezoelectric transducers". *Journal of electroceramics*, 19(1):167–184 (2007).
41. Zabek, D., Taylor, J., Ayel, V., Bertin, Y., Romestant, C., and Bowen, C. "A novel pyroelectric generator utilising naturally driven temperature fluctuations from oscillating heat pipes for waste heat recovery and thermal energy harvesting". *Journal of Applied Physics*, 120(2) (2016).
42. Chiolerio, A., Garofalo, E., Mattiussi, F., Crepaldi, M., Fortunato, G., and Iovieno, M. "Waste heat to power conversion by means of thermomagnetic hydrodynamic energy harvester". *Applied Energy*, 277:115591 (2020).
43. Suo, G., Yu, Y., Zhang, Z., Wang, S., Zhao, P., Li, J., and Wang, X. "Piezoelectric and triboelectric dual effects in mechanical-energy harvesting using batio3/polydimethylsiloxane composite film". *ACS applied materials & interfaces*, 8(50):34335–34341 (2016).
44. Lang, S.B. "Pyroelectricity: from ancient curiosity to modern imaging tool". *Physics today*, 58(8):31–36 (2005).
45. Bowen, C.R., Taylor, J., LeBoulbar, E., Zabek, D., Chauhan, A., and Vaish, R. "Pyroelectric materials and devices for energy harvesting applications". *Energy & Environmental Science*, 7(12):3836–3856 (2014).
46. Barbier, E. "Geothermal energy technology and current status: an overview". *Renewable and sustainable energy reviews*, 6(1-2):3–65 (2002).
47. Avirovik, D., Kumar, A., Bodnar, R.J., and Priya, S. "Remote light energy harvesting and actuation using shape memory alloy—piezoelectric hybrid transducer". *Smart Materials and Structures*, 22(5):052001 (2013).
48. Avirovik, D., Kishore, R.A., Vuckovic, D., and Priya, S. "Miniature shape memory alloy heat engine for powering wireless sensor nodes". *Energy Harvesting and Systems*, 1(1-2):13–18 (2014).
49. Bai, S. and Liu, C. "Overview of energy harvesting and emission reduction technologies in hybrid electric vehicles". *Renewable and Sustainable Energy Reviews*, 147:111188 (2021).
50. Aksu, U. and Halicioglu, R. "A review study on energy harvesting systems for vehicles". *Tehnički glasnik*, 12(4):251–259 (2018).
51. Li, D., Wu, Y., Da Ronch, A., and Xiang, J. "Energy harvesting by means of flow-induced vibrations on aerospace vehicles". *Progress in Aerospace Sciences*, 86:28–62 (2016).

52. Hewawasam, L., Jayasena, A., Afnan, M., Ranasinghe, R., and Wijewardane, M. "A review paper on use of thermo-electric materials for future energy harvesting in automobile sector". In *2018 8th International Conference on Power and Energy Systems (ICPES)*, pages 157–160. IEEE (2018).
53. Maurya, D., Kumar, P., Khaleghian, S., Sriramdas, R., Kang, M.G., Kishore, R.A., Kumar, V., Song, H.C., Park, J.M.J., Taheri, S., et al. "Energy harvesting and strain sensing in smart tire for next generation autonomous vehicles". *Applied energy*, 232:312–322 (2018).
54. Anton, S.R. and Sodano, H.A. "A review of power harvesting using piezoelectric materials (2003–2006)". *Smart materials and Structures*, 16(3):R1 (2007).
55. Malladi, V.V.S., Albakri, M., Musgrave, P., and Tarazaga, P.A. "Investigation of propulsive characteristics due to traveling waves in continuous finite media". In *Bioinspiration, Biomimetics, and Bioreplication 2017*, volume 10162, pages 112–119. SPIE (2017).
56. Musgrave, P.F., Albakri, M.I., Tenney, C., and Tarazaga, P.A. "Generating and tailoring structure-borne traveling waves on two-dimensional surfaces". *Journal of Sound and Vibration*, 480:115417 (2020).
57. Sfakiotakis, M., Lane, D.M., and Davies, J.B.C. "Review of fish swimming modes for aquatic locomotion". *IEEE Journal of oceanic engineering*, 24(2):237–252 (2002).
58. Anakok, I., Malladi, V.S., and Tarazaga, P.A. "A theoretical study on the generation and propagation of traveling waves in strings". In *Topics in Modal Analysis & Testing, Volume 8: Proceedings of the 37th IMAC, A Conference and Exposition on Structural Dynamics 2019*, pages 311–316. Springer (2019).
59. Phoenix, A., Malladi, V.S., and Tarazaga, P.A. "Traveling wave phenomenon through piezoelectric actuation of a free-free cylindrical tube". In *Smart Materials, Adaptive Structures and Intelligent Systems*, volume 57304, page V002T04A018. American Society of Mechanical Engineers (2015).
60. Anakok, I., Malladi, V.S., and Tarazaga, P.A. "A study on the generation and propagation of traveling waves in strings". In *Topics in Modal Analysis & Testing, Volume 9: Proceedings of the 36th IMAC, A Conference and Exposition on Structural Dynamics 2018*, pages 257–261. Springer (2018).
61. Anakok, I., Davaria, S., Tarazaga, P.A., and Malladi, V.V.S. "A study on steady-state traveling waves in one-dimensional non-dispersive finite media". *Journal of Sound and Vibration*, 528:116907 (2022).
62. Avirovik, D., Malladi, V.S., Priya, S., and Tarazaga, P.A. "Theoretical and experimental correlation of mechanical wave formation on beams". *Journal of Intelligent Material Systems and Structures*, 27(14):1939–1948 (2016).
63. Malladi, V.V.N.S. "Continual traveling waves in finite structures: theory, simulations, and experiments" (2016).
64. Malladi, V.S., Avirovik, D., Priya, S., and Tarazaga, P.A. "Travelling wave phenomenon through a piezoelectric actuation on a free-free beam". In *Smart Materials, Adaptive structures and intelligent systems*, volume 46148, page V001T03A017. American Society of Mechanical Engineers (2014).
65. Malladi, V., Avirovik, D., Priya, S., and Tarazaga, P. "Characterization and representation of mechanical waves generated in piezo-electric augmented beams". *Smart Materials and Structures*, 24(10):105026 (2015).
66. Malladi, V.S., Albakri, M., Tarazaga, P.A., and Gugercin, S. "Reduced plate model used for 2d traveling wave propagation". In *Smart Materials, Adaptive Structures and Intelligent Systems*, volume 57298, page V001T03A021. American Society of Mechanical Engineers (2015).
67. Malladi, V.V.S., Albakri, M., and Tarazaga, P.A. "An experimental and theoretical study of two-dimensional traveling waves in plates". *Journal of Intelligent Material Systems and Structures*, 28(13):1803–1815 (2017).
68. Musgrave, P.F., Sriram Malladi, V., and Tarazaga, P.A. "Generation of traveling waves in a 2d plate for future drag reduction manipulation". In *Special Topics in Structural Dynamics, Volume 6: Proceedings of the 34th IMAC, A Conference and Exposition on Structural Dynamics 2016*, pages 129–138. Springer (2016).
69. Musgrave, P.F. and Tarazaga, P.A. "Turbulent boundary layer over a piezoelectrically excited traveling wave surface". In *AIAA Scitech 2019 Forum*, page 1354 (2019).
70. Krishnan, M., Malladi, V.V.S., and Tarazaga, P.A. "Leveraging a data-driven approach to simulate and experimentally validate a mimo multiphysics vibroacoustic system". *Mechanical Systems and Signal Processing*, 166:108414 (2022).
71. Davaria, S., Krishnan, M., Sriram Malladi, V.V., and Tarazaga, P.A. "Miniature underwater robot–an experimental case study". In *Special Topics in Structural Dynamics & Experimental Techniques, Volume 5: Proceedings of the 40th IMAC, A Conference and Exposition on Structural Dynamics 2022*, pages 119–122. Springer (2022).
72. Avitabile, P. *Modal testing: a practitioner's guide*. John Wiley & Sons (2017).
73. Bucher, I. "Estimating the ratio between travelling and standing vibration waves under non-stationary conditions". *Journal of sound and vibration*, 270(1-2):341–359 (2004).
74. Gadekar, S.S. "An experimental study towards underwater propulsion system using structure borne traveling waves". Master's thesis, Michigan Technological University (2022).
75. Gupta, S. and Malladi, V.V.S. "Utilizing steady-state traveling waves in a quiescent water environment for particle propulsion". In *Smart Materials, Adaptive Structures and Intelligent Systems*, volume 88322, page V001T03A008. American Society of Mechanical Engineers (2024).

Chapter 13
Vision-Based State Estimation of a Serial Manipulator

Aishani Pathak, James Halverson, Kenny DeCay Jr., Anand Iyer, Joshua Tempelman, Cole Maxwell, and Ricardo Mejia-Alvarez

Abstract We present an approach for estimating the configuration of a robotic manipulator using computer vision. While sensors and encoders are built into the design of most industrial robotic arms, there are settings where integrating such components is infeasible due to constraints such as cost, complexity, or operational environment. In such cases, alternative methods of state estimation are required. This work leverages a transfer learning-based method to extract the joint configuration of a robotic arm in 3D space from images. Using a pretrained vision model as a backbone, we tune the network to extract informative feature maps corresponding to each degree of freedom (DOF) of the manipulator; a dense network subsequently infers the corresponding state vector. We propose a stage-based approach with a simplified model training pipeline to estimate the state of a serial manipulator within a real environment.

Keywords Vision · State Estimation · Manipulator · Model · Encoderless

Introduction

State estimation is the process of inferring the current state (e.g. position, velocity, orientation) of a dynamic system. In robotics, this allows for real-time localization where the configuration of an arm is determined through the estimation of the angular and translational components of individual joints or linkages[1, 2]. Most robotic systems accomplish this using embedded encoders, with each joint that composes an arm contains an encoder which returns its current state [3]. This state measurement is then traditionally used as the feedback element for a control system. **Figure 1** depicts a generic control system. Error is defined as the difference between the reference input angle and the output angle. A negligible error indicates completion of the system translation [4].

Aishani Pathak
Arizona State University, Tempe, Arizona, U.S.
e-mail: ajpathak@asu.edu

James Halverson
Michigan Technological University, Houghton, Michigan, U.S.
e-mail: jhalvers@mtu.edu

Kenny DeCay Jr.
Texas A&M University, College Station, Texas, U.S.
e-mail: kdecay@tamu.edu

Anand Iyer · Joshua Tempelman · Cole Maxwell
Los Alamos National Laboratory, Los Alamos, New Mexico, U.S.
e-mail: ani.iyer@lanl.gov; tempelman@lanl.gov; colemaxwell@lanl.gov

Ricardo Mejia-Alvarez
Michigan State University, East Lansing, Michigan, U.S
e-mail: rimejal@egr.msu.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

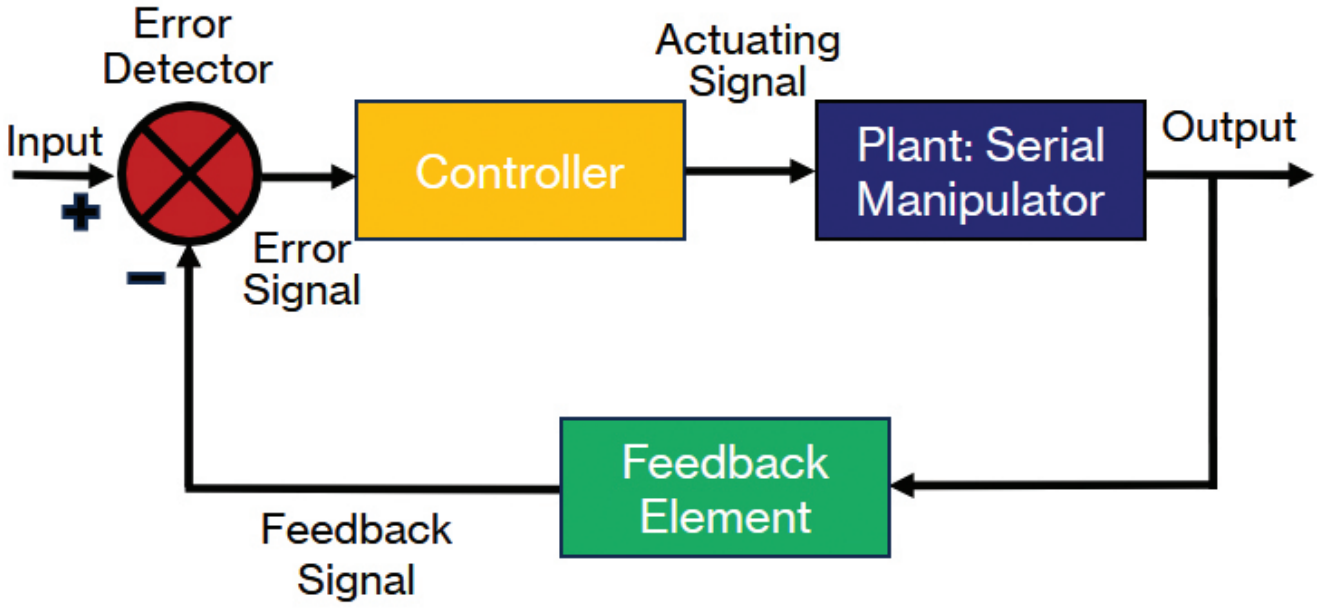


Fig. 1 Basic closed-loop control system diagram.

Although encoder based state estimation is robust and has proven precision, there are numerous scenarios in which encoder reliability struggles or are infeasible to implement. For example, in high radiation environments where gamma and beta radiation interfere with encoders [5], or on-orbit servicing operations where encoder failure can lead to neighboring object damage [6] and replacement costs are prohibitive. These constraints motivate the need for alternative state estimation methods. One of the predominant alternatives is the use of vision for encoderless state estimation and control. In general, previous vision based approaches can be separated into two categories: two-dimensional (2D) and three-dimensional (3D) methods. 3D methods primarily involve the use of RGB-Depth cameras or Light Detection and Ranging (LiDAR) systems [7, 8]—these have been effective for mobile robot object tracking or end effector state estimation, but often suffer from self occlusion and high computational expense. Further information on 3D methods please reference [9, 10, 11] To avoid the limitations of 3D vision methods, this considers a 2D method. In contrast to 3D methods, 2D methods utilize traditional imagery, making them less computationally expensive, as they perform feature extraction solely in the image plane. This computational savings is essential for vision based control, where near real-time feedback is essential. There are two main 2D image based techniques for state estimation. The first is the use of fiducial markers where physical points or markings are placed onto the system. These predefined points serve as tracking markers to determine the current state. The major drawback of this method is that the accuracy of detecting these markers suffers under camera motion blur, self-occlusion when markers are blocked from the camera, and false-positive detections from camera artifacts [12]. The second is the use of keypoints, whereby non-physical points are algorithmically placed on to features that can be reliably tracked for state estimation. Although the use of keypoints is often more robust, the detection accuracy of these virtual markers can also suffer from self-occlusion, false feature matches, and environmental lighting conditions. In addition, this method is also far more computationally expensive [13].

The primary drawback of the methods described above is their lack of generalizability. The 2D methods described above are largely intended for use on a single manipulator, and do not inherently generalize well to different manipulator configurations or alternate robotic arms. Recently, advancements in machine learning have allowed for more robust robotic state estimation using 2D imagery. Machine learning, and more specifically computer vision, has shown promise in bridging this generalization gap. In this paper, we present an alternative 2D method of state estimation using machine learning based computer vision, through the implementation of a convolutional neural network (CNN) trained on planar images of a serial manipulator. Through this approach we demonstrate the ability of a CNN architecture to perform state estimation on a serial manipulator without modification to the arm itself, using purely vision. This methodology allows for future adaption of this model to alternate arms or configurations as there is no embedded information of our specific training manipulator. This work provides motivation for further research into computer vision based inference and control methods that can provide accurate and robust state estimation without the use of encoders.

Methods

To leverage a computer vision model for state inference, large amounts of labeled data is needed for training. As such, both virtual and physical environments were created to collect data from a physical serial manipulator. To avoid the reliance on time-consuming real-world data collection, a physics-based simulation was utilized to curate a large and diverse dataset of training images. Although training with idealized data is not wholly representative of the desired physical environment, further tuning to this desired space requires much less real-world data. Specifically, we selected a transfer learning approach using ResNet-18 pretrained on ImageNet for its feature extraction abilities, with our target task being a regression over five joint angles. Finally, we considered various different training strategies to find an optimal model for this system, e.g., different loss functions, frozen ResNet parameters, shuffled batches, and weighted loss functions.



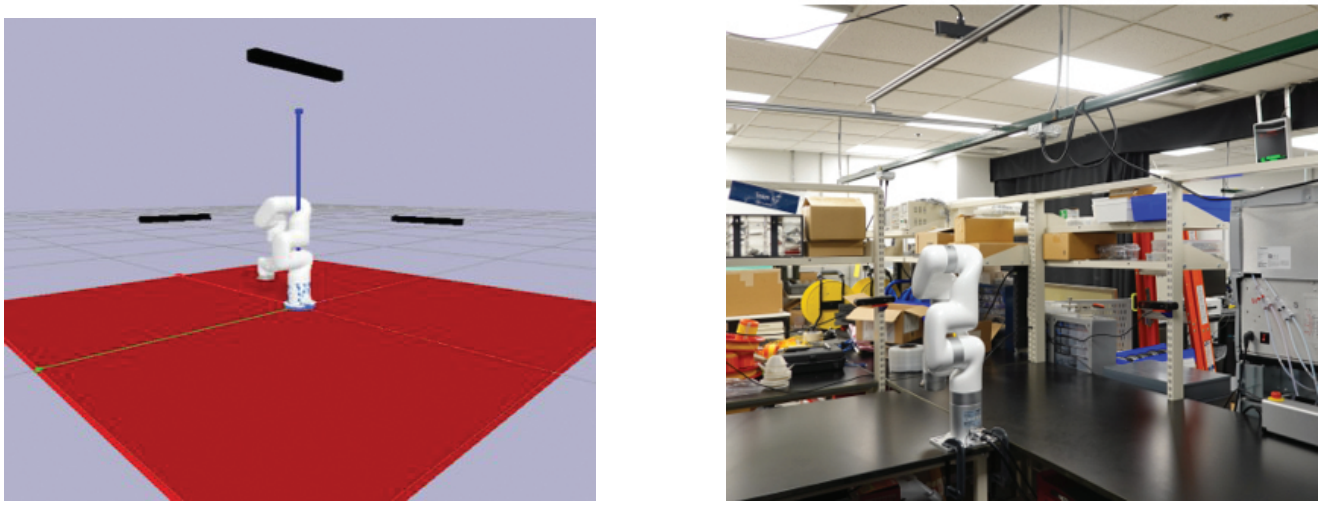
Fig. 2 Real-world setup with cameras marked.

Real-World Setup

The physical environment and experimental setup collection of real-world data were carried out as follows. A monochrome UFACTORY X-ARM 6 serial manipulator was mounted to a table and observed with three RGB cameras positioned in orthogonal directions. **Figure 2** depicts the setup and three orthogonal chosen to optimize the independence of each image. Tarps were hung in the background of the manipulator with respect to each camera view to reduce distractions present in the physical environment. While this strategy reduced clutter, natural imperfections such as lighting variations and reflective surfaces remained to ensure that the physical data retained realistic challenges. Experimental images were captured using three ZED 2i RGB cameras with a 70° vertical field of view (FOV) and an aspect ratio of 16:9.

Simulation Setup

The simulation environment was implemented using PyBullet, which is a Python-based interface to the Bullet Physics Engine. PyBullet provides a lightweight simulation framework with the ability to support various robot models and perform collision detection and camera rendering. The simulation environment was configured to emulate the physical setup, replicating the manipulator and the three camera configuration. **Figure 3** shows the manipulator in both the simulated and physical environments, with the arm configuration being positioned identically and captured from the same camera angle. Moreover, the simulation also incorporated a uniform black background to maintain further consistence with the tarp background of the physical environment (not shown).



(a) Simulation environment (b) Real environment

Fig. 3 Comparison between (a) simulated and (b) real environments with identical manipulator states and viewpoints.

This ensured geometric and visual consistency between simulation and physical environments allowing exploration of a wide range of manipulator configurations at a low computational cost. A benefit of the simulation framework is its ability to automatically flag invalid configurations, i.e., a set of joint angles would cause the manipulator to collide with itself or with a rigid object in its environment. Using the "getContactPoints" method, which detects when collidable geometry intersect, these configurations were excluded to better represent reality and to reduce the size of the training set. Moreover, to ensure physical consistency of the PyBullet simualtions, several verification and validation strategies were employed including MATLAB Simulink, kinematic calculations, and the physical manipulator's software package.

Simulation images were captured using the same FOV utilized by the cameras in the physical setup. For reference, the three camera views at 40° and 70° vertical FOV are shown in **Figure 4**. Using an increased FOV with a constrained image size does decrease the angular resolution of the image, seen by the stretched or otherwise distorted geometry near the edges of the bottom row of images [14]. However, given the nature of a regression model and our method of preprocessing, we do not expect this smearing to negatively impact performance.

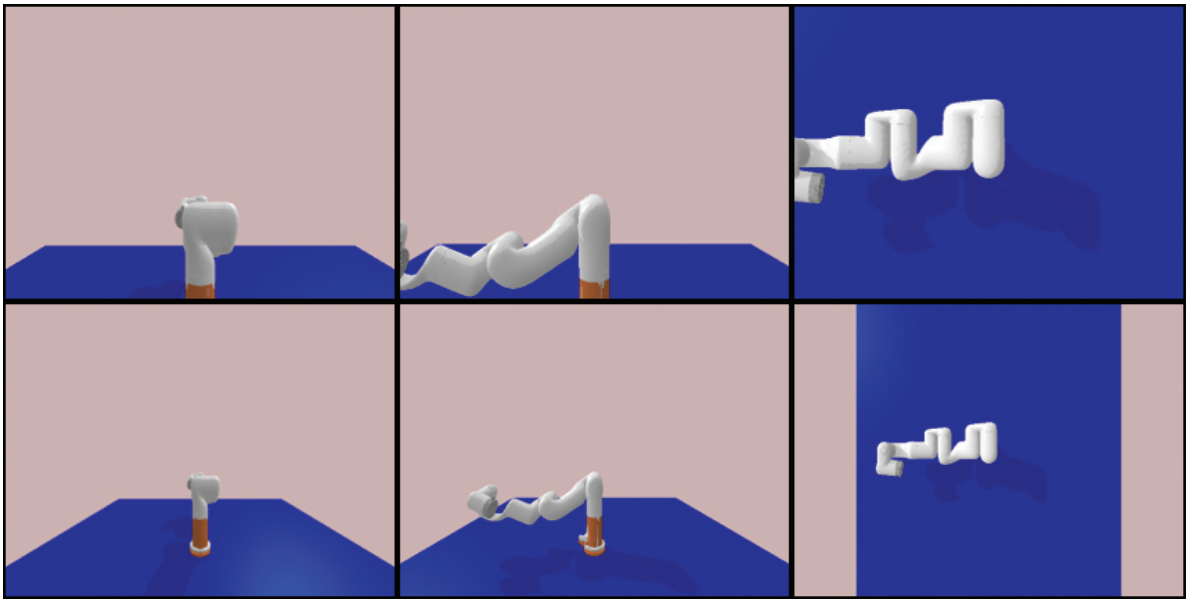


Fig. 4 Orthogonal camera perspectives for 40° vertical FOV (first row) and 70° vertical FOV (second row), as seen in PyBullet.

Data Curation and Collection

Prior to data acquisition, the valid configuration were first determined. We comment that only the first five joints of the UFACTORY X-Arm 6 were considered. In the absence of an end effector, the final wrist joint was excluded due to the small physical size of the final link.

Using the representative PyBullet model, joint angles were discretized at 15° increments, providing sufficient coverage of the configuration space for the model to interpolate between states. Since sampling the entire range of -2π to 2π led to duplicate configurations, the limits for these joints were constrained. Additionally, a small offset of 0.1 radians was introduced instead of a hard zero to avoid overlap with the upper bound at 2π, where the joint yet again assumed the same physical orientation as at 0 radians. These adjustments are highlighted in **Table 1**.

Table 1 Comparison of the manufacturers joint limitations and the modified joint limits in units of radians for the first five joints.

Joint Number	Original		Modified	
	Lower	Upper	Lower	Upper
1	-6.283	6.283	0.1	6.283
2	-2.059	2.094	-2.059	2.094
3	-3.927	0.192	-3.927	0.192
4	-6.283	6.283	0.1	6.283
5	-1.693	3.142	-1.693	3.142

Data curation resulted in a total of 75,832 possible configurations. Nearly all these images were collected from the simulation environment and later used to train the ResNet-18 model to the simulation space. The dataset was augmented with 500 samples from the physical environment which were later used to tune the computer vision to real-world measurements.

Despite prior efforts to exclude invalid configurations from this set, certain configurations did not account for the path from one configuration to the next and were thus not achievable in the real-world environment. Additionally, the self-collision detection built into the physical manipulator's controller was more sensitive than the simulation's "getContact-Points" method, meaning that the controller would prevent the manipulator from achieving it. Efforts were taken to address these challenges by computing the minimum distances to collidable geometries and imposing a threshold. However, it was deemed most practical to manually reject colliding configurations from the relatively small data set.

Preprocessing

Once data was collected over both environments, training and testing datasets were preprocessed to provide standardized input to the model. Given that the manipulator only occupies a portion of each image, each orthogonal view was first center-cropped to 470×470 pixels. This crop size was chosen to retain regions relevant to the manipulator's motion while discarding less informative background areas. The cropping not only preserves important features but reduces the computational load as well by limiting the number of processed pixels. Following this, images were resized to 200×200 to match the input dimensions required by ResNet-18 and was normalized using ImageNet's expected channel-wise mean and standard deviation: mean = [0.485, 0.456, 0.406], std = [0.229, 0.224, 0.225]. Finally, the images were concatenated horizontally producing a single 200×600 image to form a composite representation of the manipulator from all three directions.

Alongside image preprocessing, target joint angles were encoded in sine-cosine form as $(\sin\theta, \cos\theta)$ to improve learning stability. Because raw angular representations can create discontinuities in the dataset, angles near 0 and 2π are visually almost identical but numerically appear far apart. These discontinuities can confuse a regression model, leading to poor generalization and instability in predictions. Hence, this transformation embeds angular values on the unit circle, where equivalent orientations were mapped close together providing an accurate representation space. Moreover, it improved stability during training by providing smooth target space with a bounded range $[-1, 1]$.

Model Architecture

The primary objective was to estimate the joint states of the manipulator using multi-view RGB images. To leverage existing visual representations, a transfer learning approach was employed by using a pretrained ResNet-18 model as the network backbone. The ResNet-18 model employed in this work was previously trained on ImageNet, a dataset of natural images, which provided rich feature extraction capabilities. This reduced computational cost and accelerated convergence, while still allowing the network to adapt to high level visual representations of the manipulator.

To better equip the network for the regression task, modifications were to made to the final layer of the network architecture. The single standard fully connected layer of the pretrained ResNet-18 was replaced with a regression head. This layer was implemented as a multi-layer perceptron (MLP) consisting of five fully connected layers with ReLU activations and dropout for regularization. The final layer outputs 10 values, corresponding to sine and cosine encodings of the five target joint angles. The network input consisted of RGB images from three orthogonal viewpoints concatenated channel-wise, forming a single composite image input. A visual representation of the proposed model architecture can be seen in **Figure 5**.

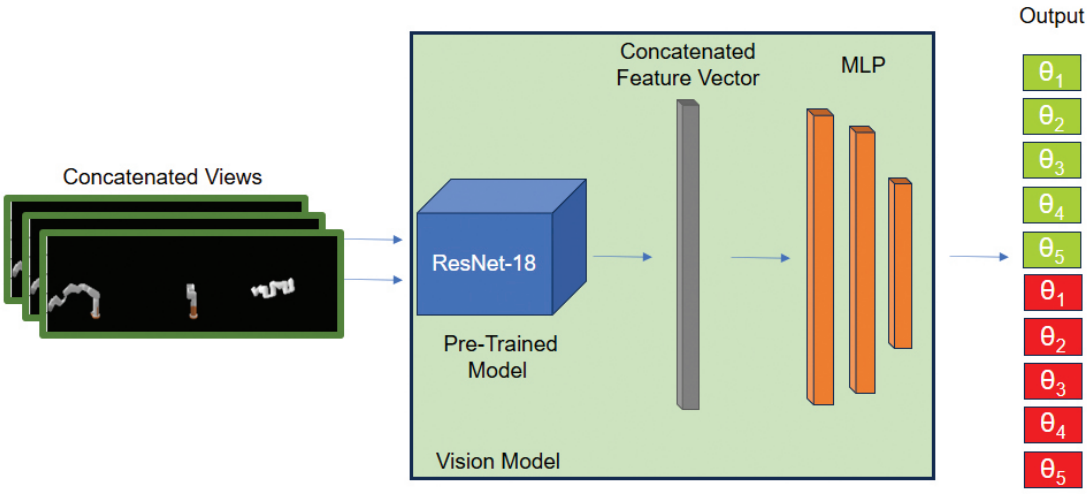


Fig. 5 Overview of the proposed vision model architecture for joint angle estimation.

Training Strategies

Models were trained using the Adam optimizer with a fixed learning rate 0.0005 and a batch size of 64. A weight decay of 1×10^{-5} was applied for the simulation-only models, while a value of 1×10^{-6} was used for the hybrid-trained models to reduce overfitting. Training was performed up to 100 epochs to test convergence, with checkpoints saved at each epoch. Several variations were applied to training to find the optimal training configuration and evaluate robustness: Two loss functions were tested: mean squared error (MSE) which penalized large deviations heavily and smooth L1 Loss (Huber Loss) which acted like piecewise function behaving quadratically for smaller error and linearly for larger errors, providing a balance between characteristics of MSE and mean absolute error (MAE). Transfer learning was applied by employing ResNet-18 parameter freezing. Two modes for the ResNet-18 backbone were: Frozen and Unfrozen with the regression head trainable across both. In addition, dataloader shuffling (True vs. False) was assessed to see its impact. The combination of these techniques yielded eight total models, which were evaluated and compared to select the best performing case.

Experiment Stages

The experimental design followed a three-stage progression to address the simulation-to-real challenge and provide subsequent results for each stage.

Stage 1, simulation-only: Models were trained, validated, and tested on synthetic data that was generated in PyBullet. This stage was essential to develop a baseline and enable training on large scale data to identify the best-performing model.

Stage 2, zero-shot transfer: The trained simulation model was then only evaluated on a set of 50 real test images. This stage measured the extent of the domain gap and assessed both environments and how closely the model generalized to real environmental data with no prior learning.

Stage 3, hybrid model: The third and final stage leveraged the converged checkpoint of the simulation model as the initialization and further retrained it on real image data. This fine-tuning was applied to leverage the knowledge of the simulation environment while incorporating real data to adapt learned representations and narrow domain gap. Evaluation was conducted on the same 50 real test images to ensure consistency between Stage 2 and Stage 3.

Results

The results are organized in the following manner: simulation-only training, zero-shot transfer to real-world data, and hybrid training with simulated and real data. Although all models were trained for 100 epochs, convergence was typically achieved within the first 20 epochs, which is selected as the point of comparison for reporting results below.

Stage 1: Simulation-only

Within the simulation-only stage, several model variants configured from our training strategies were compared to identify the best-performing model.

As seen in **Figure 6**, all models achieved similar performance, with mean absolute error falling within a narrow range. Although, it was observed that the configuration using smooth L1 loss with an unfrozen ResNet-18 using non-shuffled training (L1_UnFrozen_SF) achieved the lowest error and was therefore selected for further evaluations in real and hybrid settings. To better analyze the chosen model's performance within simulation, per-joint relative error distributions and average absolute errors are reported. Joint errors were calculated using a wrapped difference formulation. This representation mapped angular differences near 0 and 2π into the range $[-\pi, \pi]$. This wrapping was primarily beneficial for joints 1 and 4, which allow for full 2π rotation, but was applied uniformly across all joints for consistency.

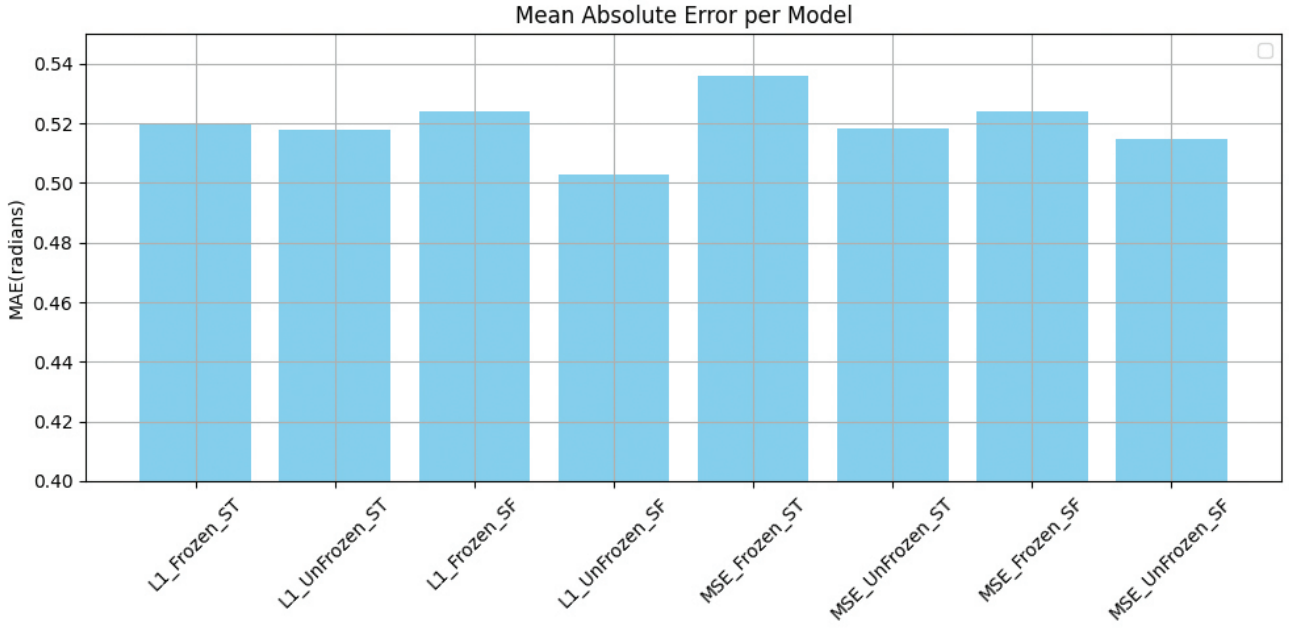


Fig. 6 Comparison of mean absolute error across all eight simulation-trained models. The y-axis is narrowed to emphasize the differences among the best-performing variants.

From the average absolute errors shown in **Figure 7** and the relative error distributions shown in **Figure 8**, it it seen that joints 1-3 have more narrowed distributions with errors concentrated around zero radians. In contrast, joints 4 and 5 have more dispersed distributions indicating higher errors. A hypothesis for this outcome is that these joints are located closed to the manipulator's end-effector, a position that is prone to self-occlusions in certain configurations and in certain cases, camera occlusions. Additionally, the links between joints 4 and 5 are shorter leading to subtler changes in appearance, which may be difficult for our network to capture consistently. Similar trends are noticed in the following evaluations as well.

Stage 2: Zero-shot Transfer

Using the best-performing simulation model case, zero-shot transfer was performed. As shown in **Figure 9**, error magnitudes increase broadly across all joints as compared to the simulation results. This is expected due to the visual domain gap: real images introduce conditions such as lighting variations and partial occlusions that are not possible to replicate in simulation training. These differences in training and testing data make it difficult for the model to extract stable visual features. The

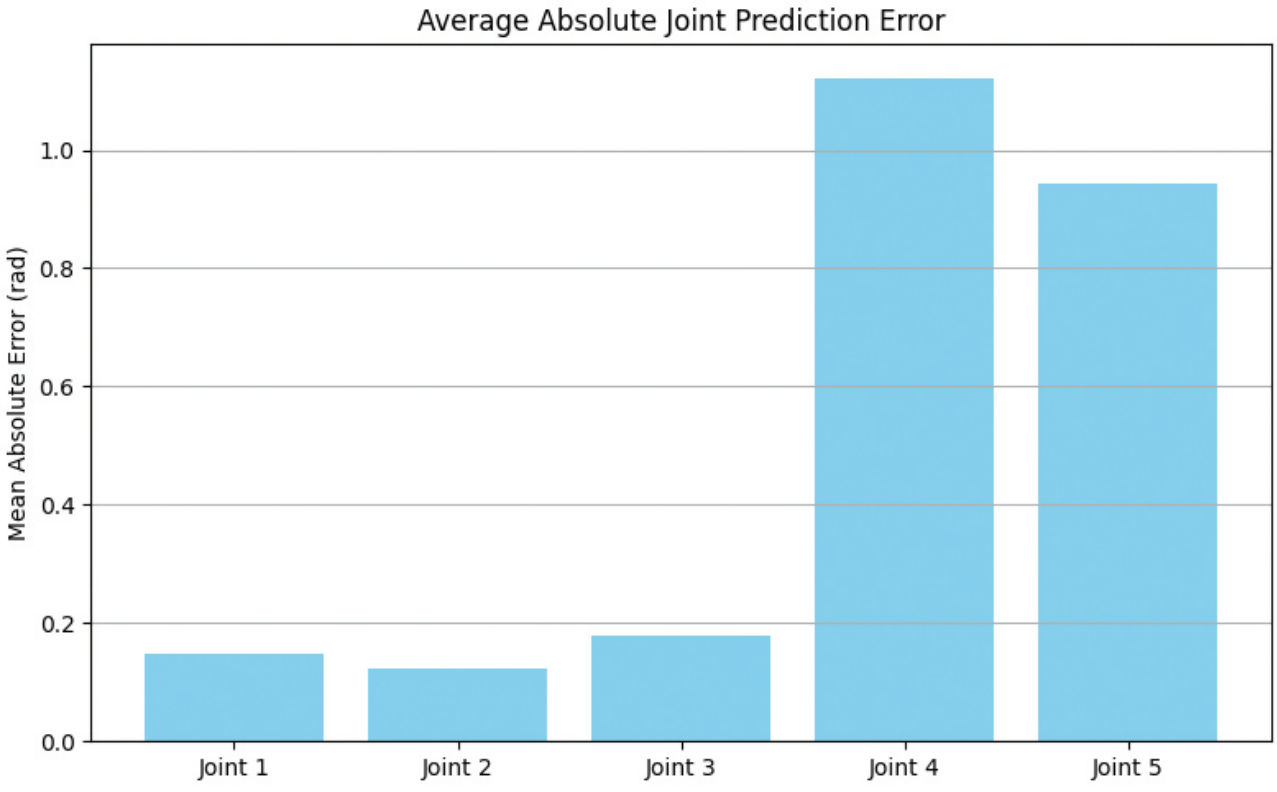


Fig. 7 Mean absolute error (MAE) for each of the five joints, averaged across simulation test samples.

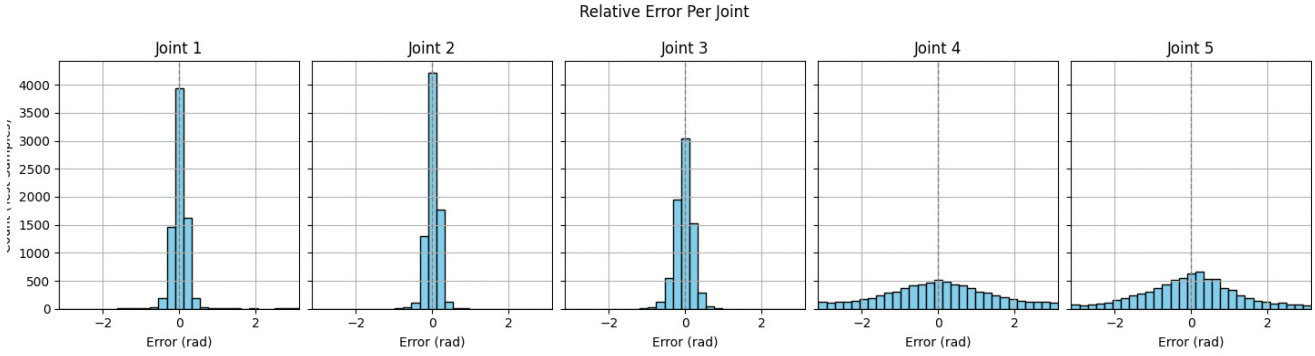


Fig. 8 Relative error distributions for each of the five joints across simulation test samples.

relative error distributions in **Figure 10** further illustrate this effect. While joint 4 and joint 5 remain the most challenging, joints 1-3 indicate that the domain gap affects the entire system rather than specific joints.

In addition to quantitative errors, a qualitative comparison of the predicted against the ground truth joint values for a single configuration is shown in **Figure 11**. Generalization is visible in the predicted results, where the overall structure of the manipulator is recovered. However, noticeable discrepancies can be supported by actual errors per joint for this case as seen in **Table 2**. Joints 1-3 and 5 suffer relatively larger errors while joint 4 stays on the lower end yet still adding to the overall error of the results. This highlights the limitation of a zero-shot transfer method from simulation to real images: the domain gap remains persistent throughout the results preventing any consistent alignment.

Finally, **Figure 12** illustrates the displacement of the end-effector for this configuration. Although the end-effector does not play a direct role in state estimation, it provides a valuable validation technique. For the zero-shot transfer case, the distance between the ground-truth and predicted end-effector is significantly large as seen by the distance line. This confirms that state estimation suffers due to domain gap.

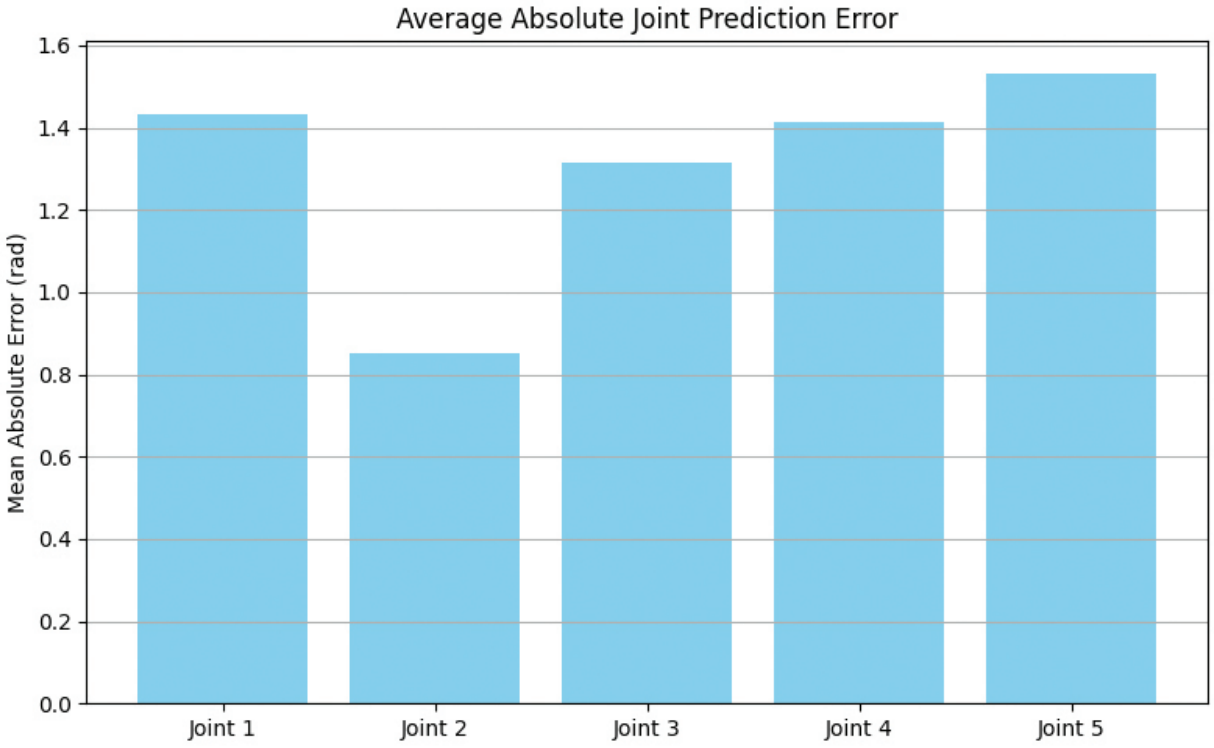


Fig. 9 Mean absolute error (MAE) per joint for the simulation-trained model evaluated on real images.

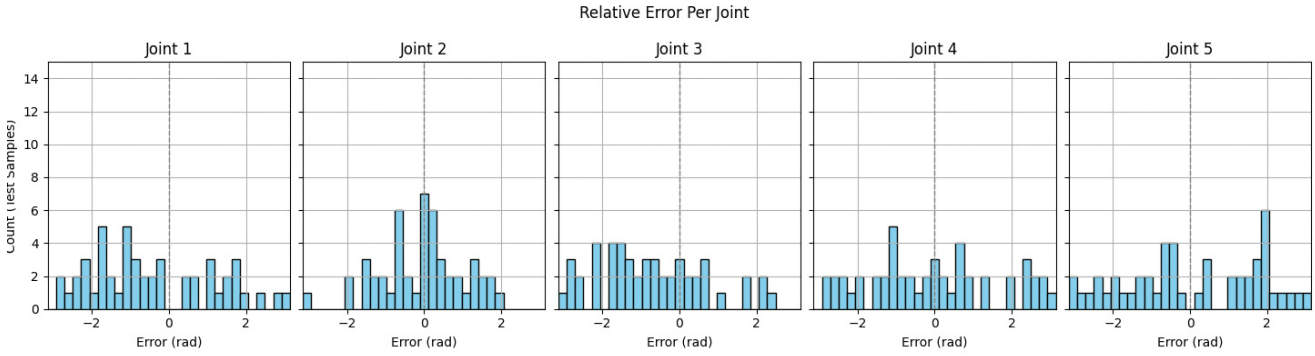
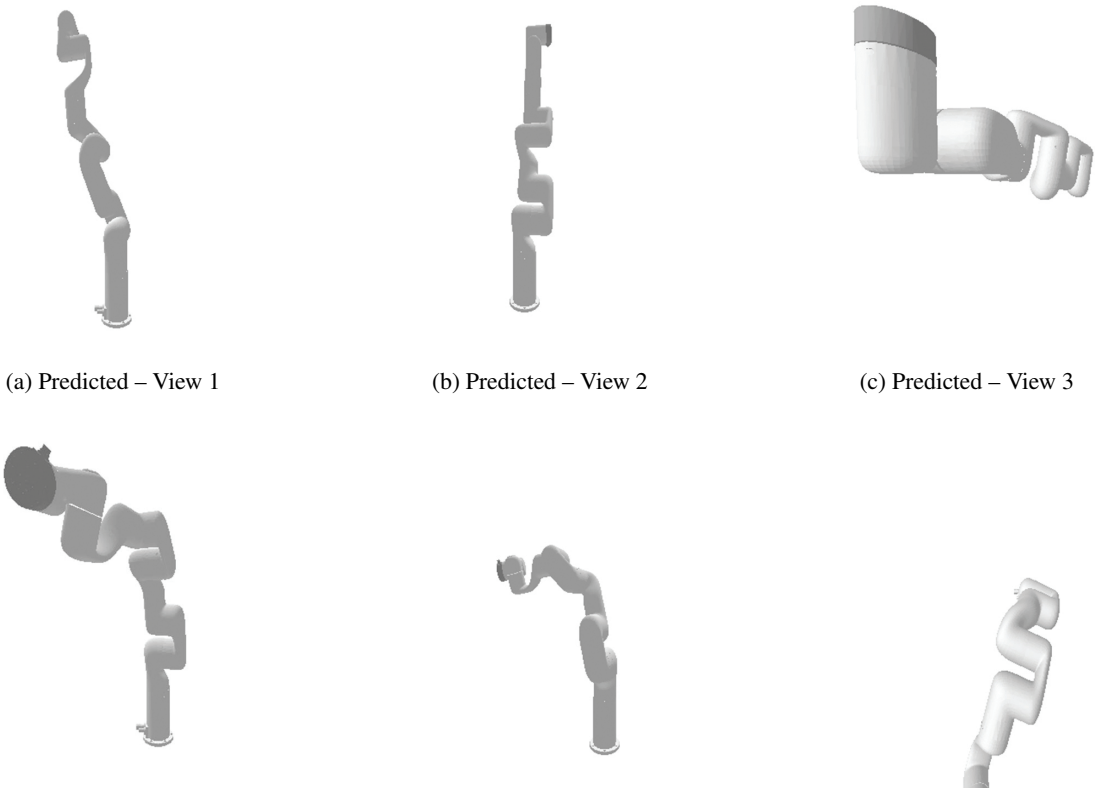


Fig. 10 Relative error distributions per joint for the simulation-trained model evaluated on real images.

Stage 3: Hybrid Model

Following the same strategy, the best performing simulation model was used for the hybrid approach.

As shown in **Figure 13**, the mean absolute error per joint decreases overall compared to the zero-shot case. Joint 1, which exhibited the worst performance in the zero-shot case, sees a substantial reduction in error and is no longer the outlier. Joint 2 also improves considerably, maintaining its position as one of the best-performing joints but with noticeably lower error than in the zero-shot evaluation. Joint 3 remains less accurate than in the pure simulation setting, yet hybrid training reduces its error relative to the zero-shot model. Joint 4 shows little change compared to zero-shot and becomes the worst-performing joint in the hybrid evaluation. Joint 5, which carried the highest error in the zero-shot case, now drops below joint 4, reflecting a modest improvement. The relative error distributions in **Figure 14** further illustrate these trends. In the zero-shot transfer case, joint 1 exhibited an extremely dispersed distribution. Under hybrid training, its errors are now tightly centered around zero indicating a significant improvement in prediction stability. Joint 2's distribution, already among the narrowest, becomes even more compact, confirming it benefits strongly from hybrid training. Joint 3 still trails



(a) Predicted – View 1 (b) Predicted – View 2 (c) Predicted – View 3

(d) Ground Truth – View 1 (e) Ground Truth – View 2 (f) Ground Truth – View 3

Fig. 11 Qualitative comparison of a real test configuration showing the predicted pose using the simulation model (first row) and the ground-truth (second row) from three views.

Table 2 Quantitative comparison on a per-joint basis of the same real test configuration, based on the simulation model.

	J1	J2	J3	J4	J5
Ground Truth (rad)	4.38	0.26	-1.75	1.22	0.45
Predicted (rad)	6.08	-0.49	-2.78	1.76	1.49
Error (rad)	**1.71**	**-0.75**	**-1.03**	**0.54**	**1.05**

its performance in simulation-only testing but shows clear improvement compared to zero-shot, with a more concentrated spread around zero. In contrast, joint 4's distribution remains largely unchanged, and joint 5 continues to be the most error-prone, though with slightly reduced tails relative to the zero-shot case. Overall, the hybrid model reduces the broad dispersion seen in the zero-shot results and produces distributions that are more consistent across joints.

To provide a direct comparison with zero-shot evaluation, **Figure 15** shows the same real test configuration rendered in PyBullet, alongside **Table 3**, which summarizes the ground-truth, predicted, and error values for each joint. Visually , the hybrid model produces a pose that is a much closer replication of the ground truth as seen in all orthogonal views. The per-joint errors in the table reinforce these visual results as they show strong numerical alignment in the manipulator joint angles. Together, the qualitative render and table confirm that hybrid training not only narrows the numerical error but also results in visibly improved alignment of the manipulator on real images.

The final evaluation process was to compare the distance between the predicted end effector location and the ground truth. As can be seen in **Figure 16**, there is error present, noted by the red line. However, this relative distance is much less than when this same comparison was made using the simulation-only model. Therefore, the performance of the model is heightened when the variability of real-world data is introduced.

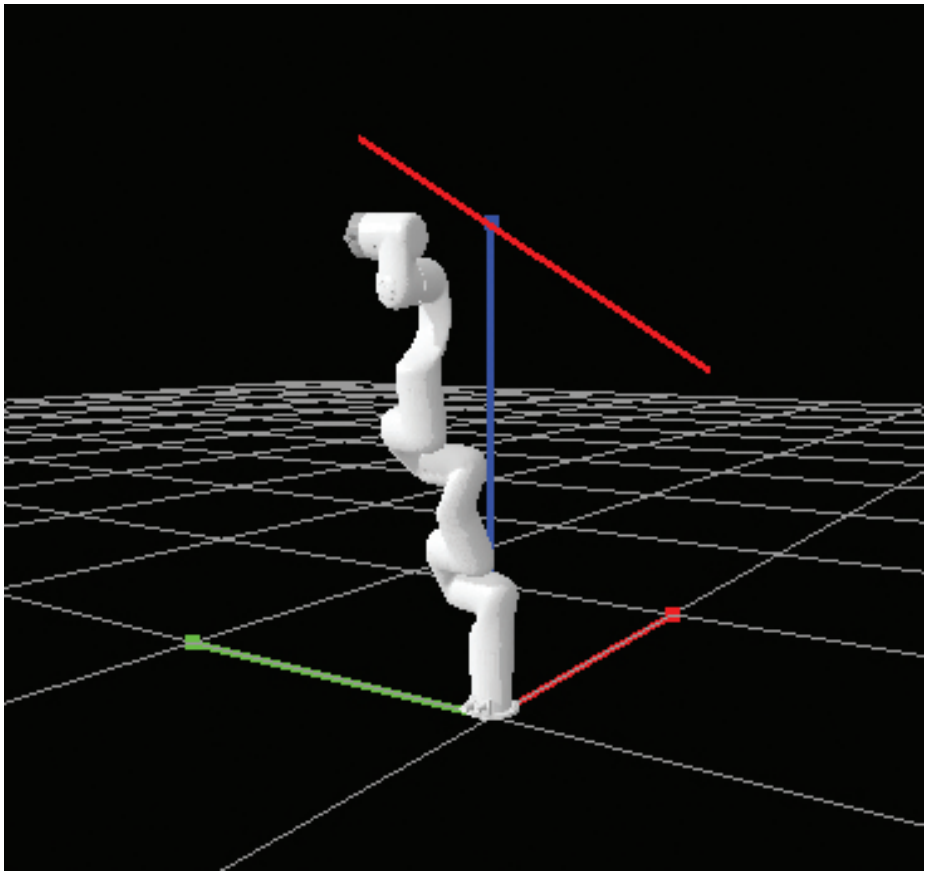


Fig. 12 Qualitative comparison of end effector position where a red line marks the Euclidean distance between the real test image and the estimated configuration using the simulation model.

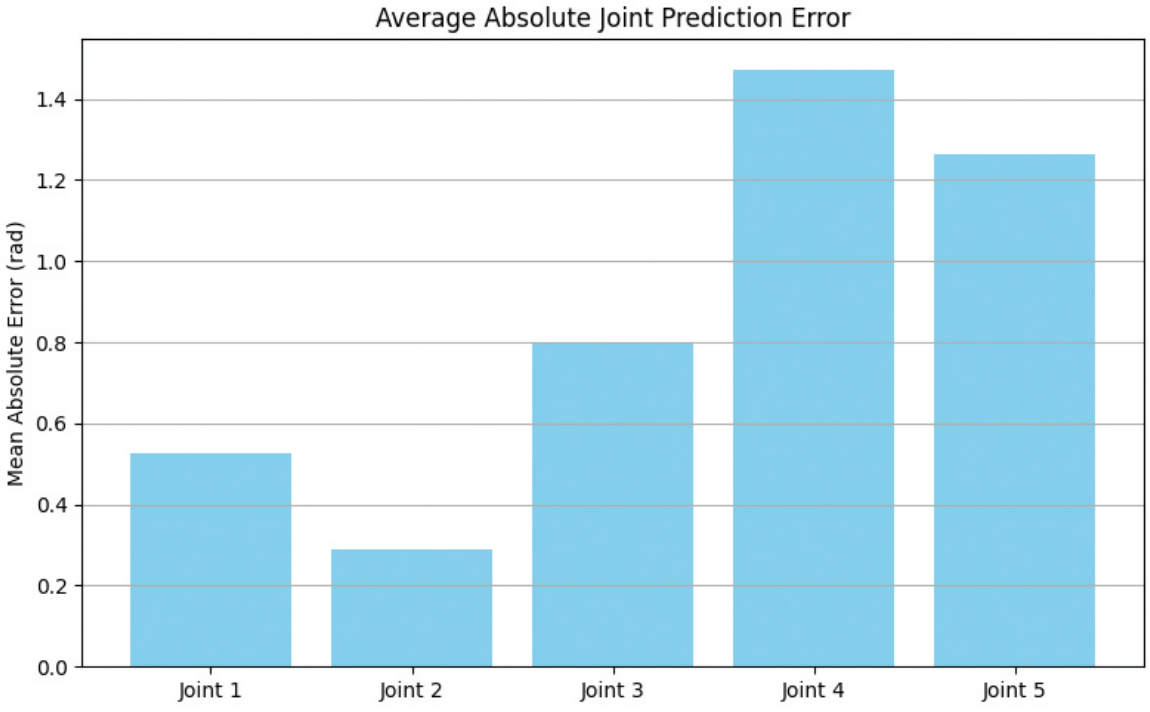


Fig. 13 Mean absolute error (MAE) per joint for the hybrid-trained model on real-image test data.

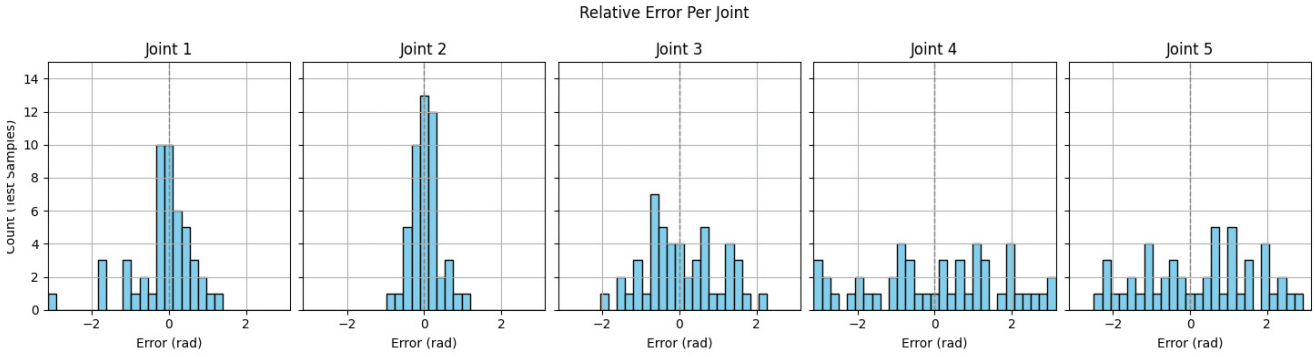
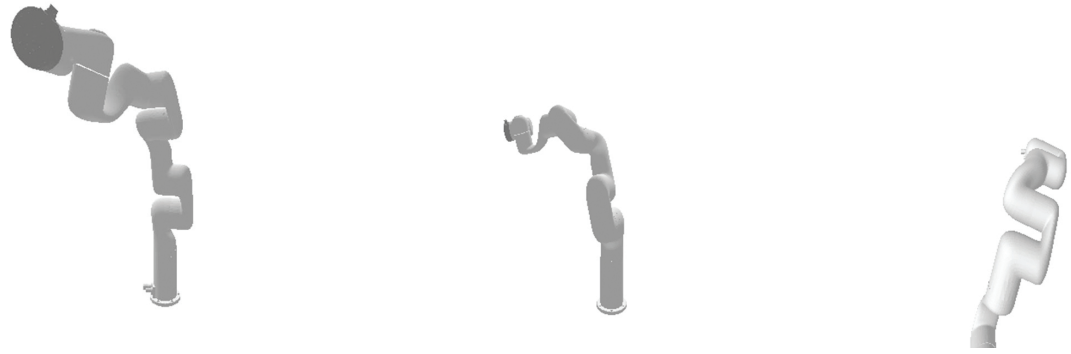


Fig. 14 Relative error distributions per joint for the hybrid-trained model on real-image test data.



(a) Predicted – View 1 (b) Predicted – View 2 (c) Predicted – View 3



(d) Ground Truth – View 1 (e) Ground Truth – View 2 (f) Ground Truth – View 3

Fig. 15 Qualitative comparison of a real test configuration showing the predicted pose using the hybrid model (first row) and the ground-truth (second row) from three views.

Table 3 Quantitative comparison on a per-joint basis of the same real test configuration, based on the hybrid model.

	J1	J2	J3	J4	J5
Ground Truth (rad)	4.38	0.26	-1.75	1.22	0.45
Predicted (rad)	4.22	-0.03	-1.56	0.54	0.52
Error (rad)	**-0.16**	**-0.29**	**0.19**	**-0.68**	**0.07**

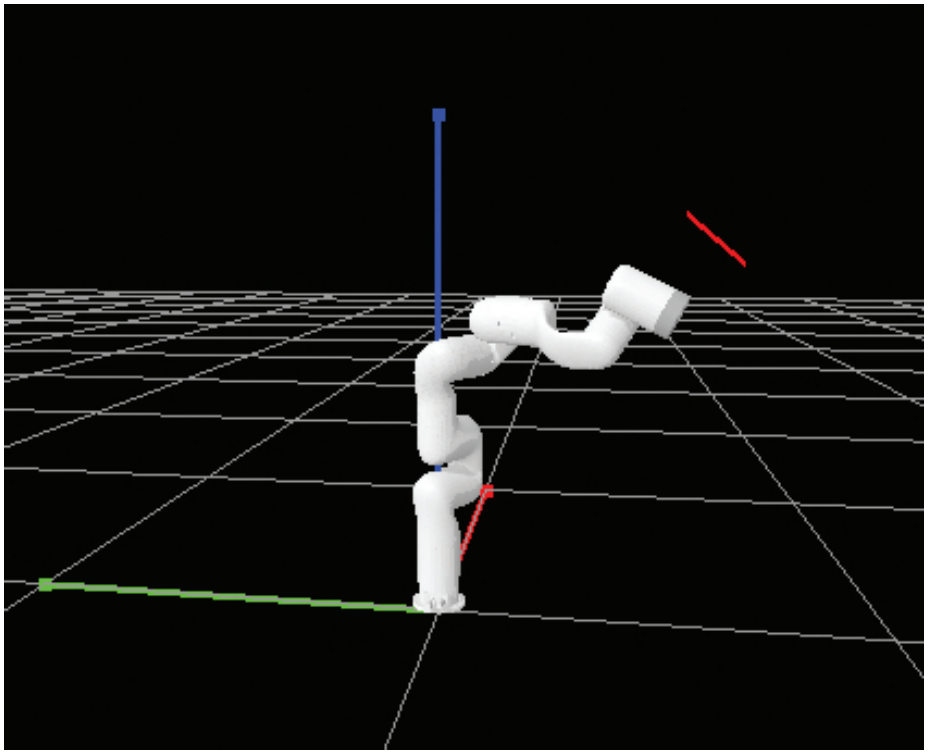


Fig. 16 Qualitative comparison of end effector position where a red line marks the Euclidean distance between the real test image and the estimated configuration using the hybrid model.

Conclusion

We developed a vision-based deep learning approach for performing state estimation of robotic joint angles from multi-view RGB inputs. Our approach incorporates a neural network with a pretrained ResNet-18 backbone, which underwent a multi-stage training process. This involved first training the network on a large set of simulated data collected from a virtual 3D environment and then fine-tuning on a smaller set of real-life camera data. The network trained using this procedure outperformed models trained exclusively on either simulated or real-life data and demonstrated reasonable performance on a set of real-life test data not seen during model training. These results validate our proposed vision-based approach for robotic state estimation and suggest promising applications for additional configurations of serial manipulators beyond those examined in this work.

Future work: This methodology can be extended in numerous directions. Inclusion of the end-effector global position as a target label for the neural network could provide additional representation to the network to adapt and learn from. This strategy could also mitigate the issue of gaps in learning for shorter links across any serial manipulator that suffer subtle visual changes across configurations. Additionally, this could help the model implicitly understand manipulator kinematics. Utilizing this vision model as an feedback signal for a serial manipulator could be a potential direction to eliminate the need for encoders, a key motivation to this work.

Broader impacts: An important contribution of this work is a baseline for encoderless manipulator functionality, which has potential applications in radiation environments, space missions, and other instances where encoder feedback is not preferred. By reducing the reliance on joint encoders, building on this approach could provide a cost-effective solution for many challenges faced in the robotics space.

Acknowledgments This research was funded by Los Alamos National Laboratory (LANL) through the Engineering Institute's Los Alamos Dynamics Summer School (LADSS). The Engineering Institute is a research and education collaboration between LANL and the University of California San Diego's Jacobs School of Engineering. This collaboration seeks to promote multidisciplinary engineering research that develops and integrates advanced predictive modeling, novel sensing systems, and new developments in information technology to address LANL mission-relevant problems.

We would also like to thank the whole of the LADSS program, and specifically Samantha Ceballes and Shannon Danforth for their excellent program organization, leadership, and mentorship.

References

1. Park, W., Liu, Y., Zhou, Y., Moses, M., and Chirikjian, G.S. "Kinematic state estimation and motion planning for stochastic nonholonomic systems using the exponential map". *Robotica*, 26(4):419–434 (2008).
2. Talbot, W., Nubert, J., Tuna, T., Cadena, C., Dümbgen, F., Tordesillas, J., Barfoot, T.D., and Hutter, M. "Continuous-time state estimation methods in robotics: A survey" (2025).
3. Sandler, B.Z. "5 - feedback sensors". In Sandler, B.Z., editor, *Robotics (Second Edition)*, pages 175–205. Academic Press, San Diego, second edition edition (1999).
4. Ceccarelli, M. *Fundamentals of mechanics of robotic manipulation*, volume 112. Springer (2022).
5. Reed, F.K., Ezell, N.D.B., Ericson, M.N., and Britton, C.L., Jr. "Radiation hardened electronics for reactor environments". Technical report, Oak Ridge National Laboratory (ORNL), Oak Ridge, TN (United States) (2020).
6. Alizadeh, M. and Zhu, Z.H. "A comprehensive survey of space robotic manipulators for on-orbit servicing". *Frontiers in Robotics and AI*, 11:1470950 (2024).
7. Han, M., Xie, B., Barczyk, M., and Bayat, A. "Image-based joint state estimation pipeline for sensorless manipulators". In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2158–2165 (2021).
8. Widmaier, F., Kappler, D., Schaal, S., and Bohg, J. "Robot arm pose estimation by pixel-wise regression of joint angles". In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 616–623 (2016).
9. Ioannidou, A., Chatzilari, E., Nikolopoulos, S., and Kompatsiaris, I. "Deep learning advances in computer vision with 3d data: A survey". *ACM Comput. Surv.*, 50(2) (2017).
10. Shirai, Y. *Three-dimensional computer vision*. Springer Science & Business Media (2012).
11. O' Mahony, N., Campbell, S., Krpalkova, L., Riordan, D., Walsh, J., Murphy, A., and Ryan, C. "Computer vision for 3d perception". In Arai, K., Kapoor, S., and Bhatia, R., editors, *Intelligent Systems and Applications*, pages 788–804, Cham (2019). Springer International Publishing.
12. Kalaitzakis, M., Cain, B., Carroll, S., Ambrosi, A., Whitehead, C., and Vitzilaios, N. "Fiducial markers for pose estimation: Overview, applications and experimental comparison of the artag, apriltag, aruco and stag markers". *Journal of Intelligent & Robotic Systems*, 101(4):71 (2021).
13. Lepetit, V. and Fua, P. "Keypoint recognition using randomized trees". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(9):1465–1479 (2006).
14. Zhang, Z., Hu, Y., Yu, G., and Dai, J. "Deeptag: A general framework for fiducial marker design and detection". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(3):2931–2944 (2023).

Chapter 14
Thickness Estimation Using Multi-Frequency Acoustic Steady-State Excitation Spatial Spectroscopy

Joseph H. Lahmann, Erica M. Jacobson, Eric B. Flynn, and Adam J. Wachtor

Abstract Acoustic steady-state excitation spatial spectroscopy is a rapid, full-field nondestructive inspection technique that excites a structure with ultrasonic transducers and measures the complex out-of-plane surface response velocity over a prescribed inspection area using a scanning laser Doppler vibrometer. Through local wavenumber estimation of the resulting spatially resolved wavefield and correlation to Lamb wave dispersion curves, local material thickness can be estimated. Because dispersion characteristics vary with excitation frequency, wavenumber, and thickness, so too does the accuracy and spatial resolution of the thickness estimates. This work seeks to improve these metrics of thickness estimation by combining steady-state measurements collected at different excitation frequencies. Full-field inspection measurements for an aluminum plate were collected using excitation frequencies ranging from 50 kHz to 500 kHz in increments of 2 kHz. Using this set of measurements, several methods for error reduction by combining scans on a pixel-level were evaluated. The best performing method fused thickness estimations informed by the full correlation between the observed and expected wavenumber spectra, and resulted in consistent error reduction with an increasing number of scans and reduced the median error from 6.5% to 2.5%.

Keywords full-field acoustic nondestructive evaluation · structural health monitoring · laser Doppler vibrometer · guided waves · wavenumber estimation

Introduction

Ultrasonic inspection techniques have proven to be a valuable tool in a wide variety of nondestructive evaluation (NDE) applications, from in-situ quality assessment of additively manufactured parts [1, 2], to the inspection and quality control of lithium-ion batteries [3]. While versatile, many ultrasonic NDE methods either lack spatial resolution or require prohibitively long inspection times to function effectively as a structural health monitoring (SHM) tool [4]. Acoustic steady-state excitation spatial spectroscopy (ASSESS) has been shown to address both of these limitations, providing rapid and spatially resolved measurements suitable for both NDE and SHM applications [5–7]. In general, ASSESS excites a structure to a steady-state condition at a single frequency and scans the surface of the structure with a laser Doppler vibrometer (LDV) to measure the surface response velocity. From this measurement, it's possible to calculate the guided wavefield [8] and then estimate the amplitude of local wavenumbers present at each pixel across the inspection space. For plate-like structures, this local wavenumber estimation method and mapping has proven useful in the estimation of coating thickness [9], detection of composite delamination [5, 10], detection and localization of corrosion [11–13], and general thickness estimation [7, 14].

Joseph H. Lahmann
Department of Mechanical Engineering, University of Cincinnati, OH 45221
Engineering Institute, Los Alamos National Laboratory, NM 87545
e-mail: lahmanjh@mail.uc.edu

Eric B. Flynn
Space Remote Sensing and Data Science, Los Alamos National Laboratory, NM 87545
e-mail: eflynn@lanl.gov

Erica M. Jacobson · Adam J. Wachtor
Engineering Institute, Los Alamos National Laboratory, NM 87545
e-mail: ejacobson@lanl.gov; ajw@lanl.gov

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

For a given excitation frequency, set of material properties, and plate thicknesses, Lamb-wave theory predicts the wavenumber and number of different modes present in the response signal [8]. Figure 1 shows an example of the expected dispersion behavior in terms of wavenumber and thickness for a 200 kHz excitation of an aluminum 6061 alloy. While Lamb-wave theory stems from infinite plates, guided wave modes will generally appear “Lamb-like" in a local region large enough to contain a few full wavelengths of the mode – which has been suggested to be around 3 wavelengths [13]. These dispersion characteristics can be leveraged to estimate local thickness of plate-like structures [14, 15]. Yet many challenges exist with thickness estimation through this framework, particularly in regard to dispersion characteristics. Firstly, there’s a wavenumber dependency on the sensitivity of the thickness estimation. For localization, a higher wavenumber is usually preferable to allow a greater number of wavelengths for a given region in order to allow the developing modes to better approximate Lamb-wave behavior. Higher wavenumbers can be achieved through increasing excitation frequency, which subsequently shifts the scale of both axes in Figure 1: increasing the magnitude of the y-axis and scaling the x-axis so that new wave modes increasingly appear at thinner thicknesses. However, the appearance of higher wave modes can complicate the thickness estimation process as the same wavenumber values can be associated with different wave modes and thicknesses. While there have been several promising deep-learning based approaches to bypass some of these nuances [16, 17], there remain fruitful lines of exploration in the analytic domain.

Several approaches have been recently explored to address some of the issues associated with the existence of multiple wave modes present in the measurements and varying sensitivity of thickness estimation. The most common approach is to evaluate modes selectively, i.e. estimate through identifying and isolating the A0 mode or the S0 mode [14]. Moon et al. formalized the idea of wavenumber sensitivity and modal separation as candidates for selecting optimal excitation frequencies, informed through comparison of FEA simulations of guided wave patterns in flat plates [13]. To take advantage of the dispersion characteristics at different frequency-thickness products in characterizing structures with a broad range of thicknesses, information from different frequencies has been fused several ways. Kang et al. built on the approach of Moon et al., selecting 3 frequencies designed to ensure the presence of a wave mode with sensitivity above an acceptable threshold across a thickness range of interest. Each wave mode identified (A0 at 200 kHz, S0 at 450 kHz and S0 at 600 kHz) was filtered out and matched to the corresponding thickness. The final thickness estimate was taken as the average across all three scans [14]. Purcell et al. implemented a chirp signal between 75 kHz to 400 kHz to utilize broadband spectral information. The A0 mode was selected via a filter bank for each thickness corresponding to the expected wavenumbers across the spectrum of frequencies. The filter that maximized the signal energy of monogenic analysis in the time-domain corresponded to the estimated thickness. It was found that broad frequency information outperformed single-frequency, resulting in a mean percent error of 14.79 % for an 8 mm aluminum test plate with varying thicknesses [15]. While demonstrated to be effective in the cases reported, these methods purge the wavenumber information from modes outside of the isolated mode rather than leveraging the additional information and incorporating it into a single estimate.

These difficulties motivate the question which was addressed in this work: *How can information from separate scans across multiple frequencies be fused to improve general thickness estimates, while accounting for the full spectrum of guided wave modes?* The subsequent sections describe the collection of a dataset ranging in excitation frequency from 50 kHz to 500 kHz, detail the calculation of a thickness estimate from the collected wavefield data, and introduce several measurement

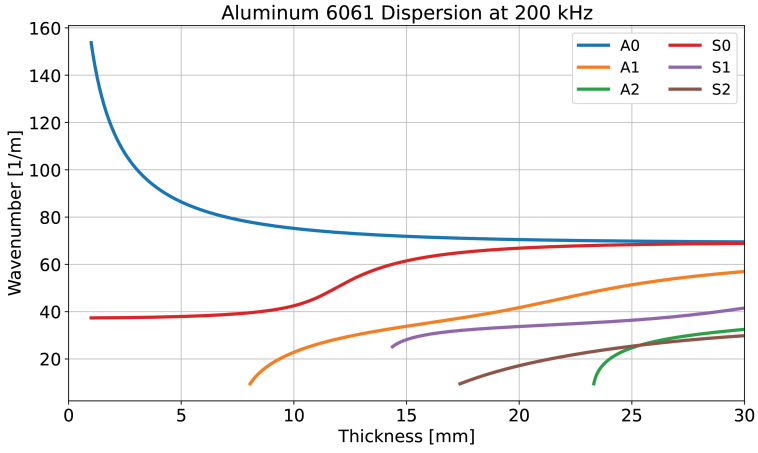


Fig. 1 Dispersion curve for aluminum 6061 at 200 kHz showing the relationship between wavenumber and structural thickness for multiple axisymmetric and symmetric modes.

fusion candidate methods for thickness estimation. These methods were compared over increasing numbers of randomly selected scan measurements to evaluate estimation error. The results indicated that fusing scans at the level of dispersion-curve similarity results in the highest potential for error reduction and defect localization.

Methods

The selected test specimen was a 10 mm aluminum 6061 plate containing a variety of manufactured thickness variations ranging from 1-10 mm in 1 mm increments, as shown in Figure 2. The diameter of the largest bores was 25 mm with varying thicknesses. The bottom row remained a constant 1 mm material thickness but decreased from 25 mm to 5 mm diameter in 5 mm increments. A transducer with rubber coupling was clamped at the bottom corner and driven with a narrow square-wave excitation at frequency, f_E, during collection. The plate was setup approximately 2 meters away from the measurement system and made level and perpendicular to the scanning head of the LDV in order to reduce distortion in the resulting wavefield. Co-registered LiDAR measurements captured the orientation of the plate normal and was used for a final data clean-up before performing wavenumber processing.

Examples of wavefield measurements from the scans that were collected from 50 kHz to 500 kHz in increments of 2 kHz are shown in Figure 3. A total of 225 independent scans were collected. One scan failed and the corrupted file was removed from the data set and also from subsequent analyses. Each scan took less than 25 seconds. With this data, a discrete vector of wavenumber and thicknesses of interest is defined, k_i and d_i respectively. Then local (pixel-wise) wavenumber was estimated by the application of a series of 2D log-Gabor filters centered at each value of k_i in the frequency-wavenumber domain,

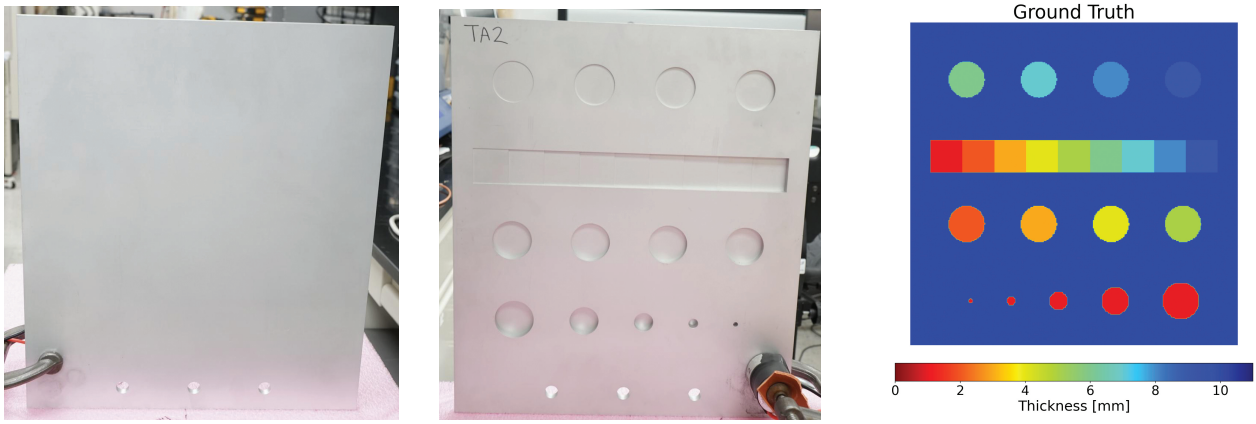


Fig. 2 Aluminum 6061 test plate containing a broad range of manufactured thickness defects. Left to right is the front (scanned side), back (thickness defects) and ground truth thickness map viewed from the front.

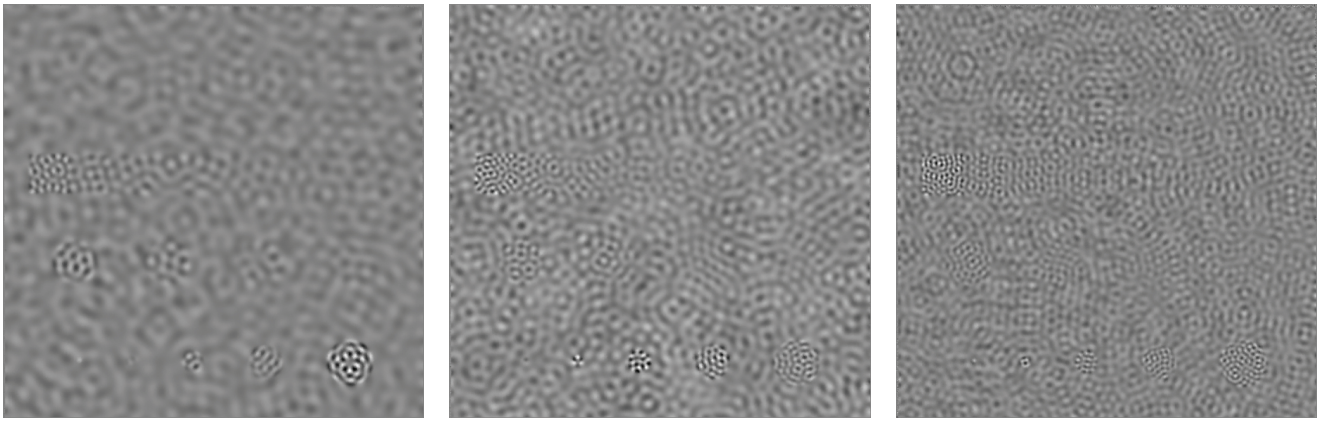


Fig. 3 Real components of complex ASSESS wavefield measurements collected at 200 kHz, 300 kHz, and 400 kHz (left to right).

inversion back to the time-space domain, and then calculation of the monogenic signal (essentially a filter-Hilbert method for complex signals) in the time domain [18]. The amplitude of the monogenic signal at each pixel for a given filter was estimated as the amplitude for that wavenumber, similar to [15]. To estimate local thickness, the resulting wavenumber spectrum was compared to the predicted thickness-wavenumber Lamb-wave dispersion curves for aluminum at each excitation frequency. This required a transformation of the wavenumber thickness dispersion curve into something more comparable.

For each mode present at thickness d, a Gaussian kernel Gabor filter was applied such that:

$$T[k,d] = \sum_m a_m[d] \cdot b_m[d] \cdot e^{\left(\frac{-(k_i - D_m[k,d])^2}{\beta(D_m[k,d]+0.001)^2}\right)} \tag{1}$$

Where T is the template, β is the filter bandwidth, $D_m[k,d]$ is the dispersion curve value for mode m at thickness d, and $b_m[d]$ is a scaling vector for higher order modes that ranges [0,1] over the first 2 mm that higher order modes are present. Furthermore, $a_m[d]$ is a second scaling factor that adjusts every mode relative to the amplitude of A0, given by:

$$a_m[d] = \frac{k_m[d]}{k_{A_0}[d]} \tag{2}$$

This results in scaling modes both according to entry and wavenumber value. Prior to use, the template was scaled to a Z-score resulting in a “simulated" spectrum seen in Figure 4.

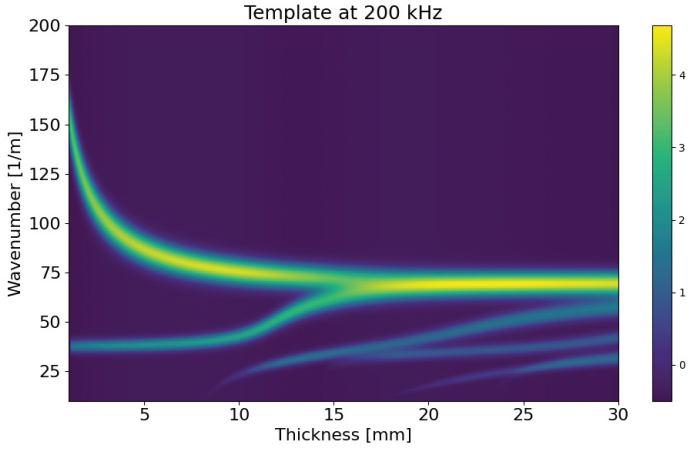


Fig. 4 A modified dispersion curve appropriate for comparison with the observed spectrum.

With this simulated spectrum, the Pearson’s correlation coefficient can be calculated between the template and the estimated wavenumber spectrum, $\hat{K}$, across all interrogated thicknesses. This resulted in a correlation curve given by:

$$c[x,y,d] = \frac{T[k,x,y,d] \cdot \hat{K}[k,x,y]}{N} \tag{3}$$

where N is the number of thicknesses interrogated. To estimate the thickness, the thickness value corresponding to the maximum correlation at each pixel was calculated. The final estimate was cleaned by a median-filter of size 7x7 pixels.

Measurement Fusion Approaches

For a given set of scans, 4 primary methods of fusion were tested to improve thickness estimates. Two naïve approaches used were the statistical mean and median, defined in Eqn. 4 and Eqn. 5.

$$\hat{d}_{\text{mean}}[x,y] = \frac{1}{N}\sum_{i=0}^{N} \hat{d}_{f_i}[x,y] \tag{4}$$

$$\hat{d}_{\text{median}}[x,y] = \text{MED}\big(\hat{d}_{f_0}[x,y], \hat{d}_{f_1}[x,y], \ldots, \hat{d}_{f_N}[x,y]\big) \tag{5}$$

However, these approaches neglect to make use of the full correlation curve. This curve can be interpreted as a measure of how similar the observed wavenumber spectrum at a pixel is to the template representing the expected Lamb-wave spectrum,

presenting information beyond the location of the maximum. One basic method of incorporating more of this information is to weight each thickness prediction with the raw amplitude of the maximum correlation, given by Eqn. 6.

$$\hat{d}_{\text{weighted mean}}[x,y] = \frac{\sum_{i=0}^{N} \max\big(c_{f_i}[x,y,d]\big) \cdot \hat{d}_{f_i}[x,y]}{\sum_{i=0}^{N} \max\big(c_{f_i}[x,y,d]\big)} \tag{6}$$

Here, c_{f_i} is the correlation between the observed wavenumber spectrum and template. A further extension would be to employ the full shape of the correlation curve, incorporating information about how similar the observed wavenumber spectrum is to every thickness. This method in general is described by Eqn. 7, which for a given set of scans will fuse each full correlation curve according to some function, prior to estimation via the maximum.

$$\hat{d}_{\text{fused correlation}}[x,y] = \arg\max_{d_i} \; f\big(c_{f_0}[x,y,d], c_{f_1}[x,y,d], \ldots, c_{f_N}[x,y,d]\big) \tag{7}$$

These methods were evaluated with sets of randomly selected frequencies. For 100 iterations, a random vector of excitation frequencies were selected and each fusion method was applied to the data with increasing numbers of scans, up to 6. Given that quality varied fairly widely in the experimental set, good performance from a fusion method should indicate robustness. For the full correlation curve fusion, only the results of a simple addition with raw correlation values are reported. While several other combinations of multiplying and scaling were explored, they resulted in nearly identical performance. Results were evaluated primarily according to absolute percent error (APE), given by:

$$\text{APE} = \left| \frac{d - \hat{d}}{d} \right| \times 100\% \tag{8}$$

which is calculated first for the entire plate. To gain a more targeted understanding of the performance, APE residuals were evaluated on the full plate, the 10 mm section of the plate (referred to as "plate"), and the defects. The defect areas were selected according to the ground truth map aligned by hand. Given the nature of the multi-scale defect plate and the sharp thickness changes from 10 to 1 mm at maximum, large absolute percent errors were expected in the residuals, resulting in a skewed distribution. Thus the summary statistics reported are the median APE (MAPE) and the 5th and 95th percentiles of the APE.

Results

The thickness calculations across all selected scans presented expected trends, *e.g.* increased spatial resolution was associated with higher values of the excitation frequency, Figure 5. At these higher frequencies, higher error was observed around the plate edge and boundaries of some of the defects. Because spatial sampling of points remains constant across all scans, this was to be expected. As shown in Figure 5, the residuals exhibited either general underestimation and overestimation.

Figure 6 shows the lowest error thickness estimate that the current single-frequency processing approach produced for the entire experimental dataset. Here the majority of thinner defects were clearly visible, but defect regions thicker than 6 mm quickly started to lose resolution, and the 9 mm thickness region and 5 mm diameter region were indistinguishable from the plate. Some neighboring frequencies resolved the smallest defect (as seen in Figure 5), but the shallow defect regions remained challenging to resolve at all frequencies.

Figure 7 compares the performance of each fusion method over successively larger sets of measurements collected at random frequencies. All the methods resulted in a decrease of the 95th percentile (worst-case performance), and a general improvement in median performance, but not necessarily the 5th percentile (best-case performance). When evaluating just the defect regions, each fusion method performed similarly, with the primary trend being a reduction in the spread of the error by increasing the number of scans. The full correlation method was the only method with a 5th percentile below a 5% residual for the defect regions and also exhibited monotonic reduction of error with increasing number of scans for all sections of the test plate. The largest improvement shown in Figure 7 was seen in the residuals of the plate. As such, the scan sets from the random selection experiment with the lowest MAPE for the plate were selected to examine in more detail.

The best thickness maps for plate estimation produced by the fusion methods are shown in Figure 8. Across all methods, the fused thickness estimates were much better at resolving the spatial features of the plate, especially for the shallower

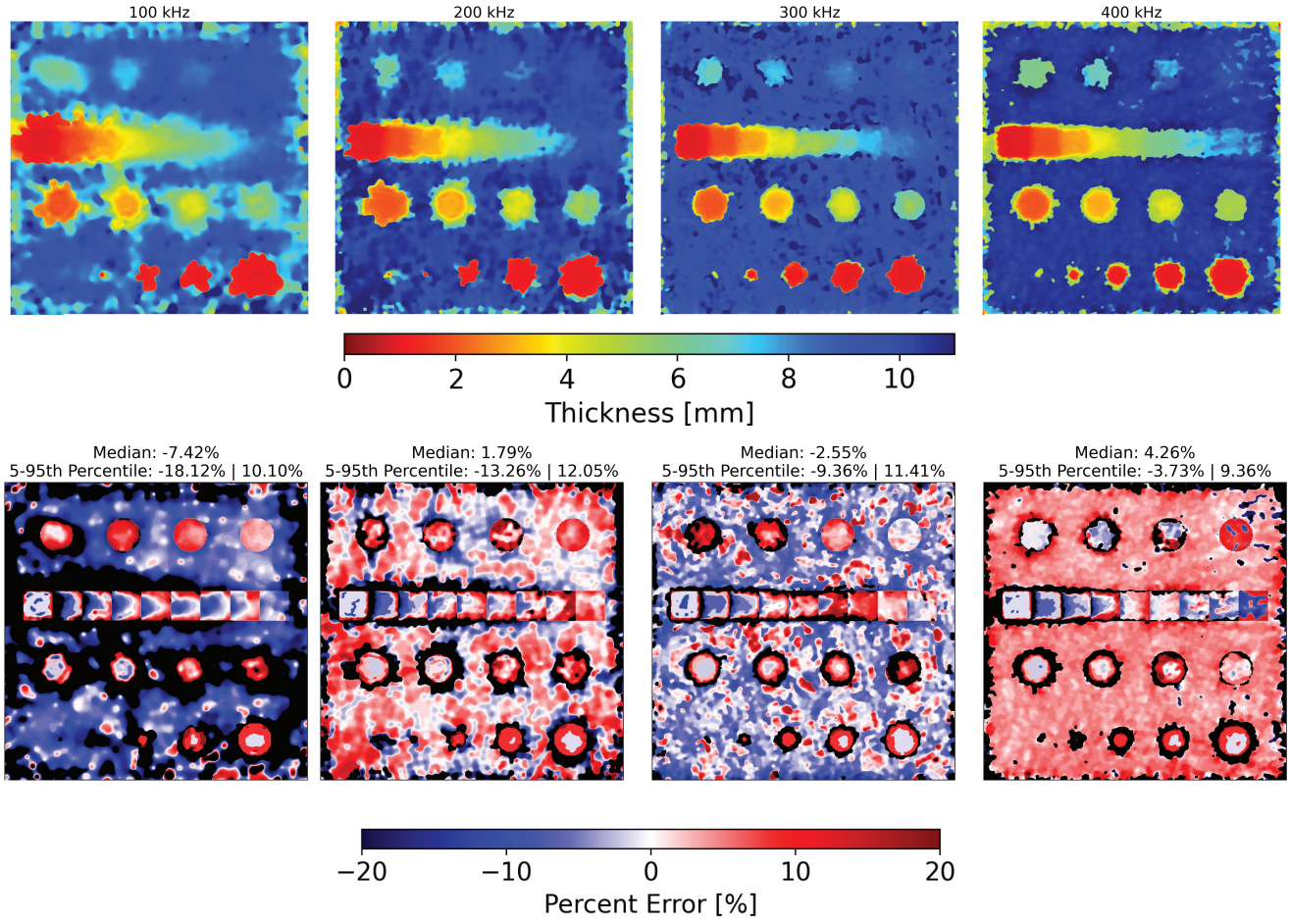


Fig. 5 Examples of single-frequency thickness estimates (top) and relative error (bottom) from 100, 200, 300, and 400 kHz excitation measurements

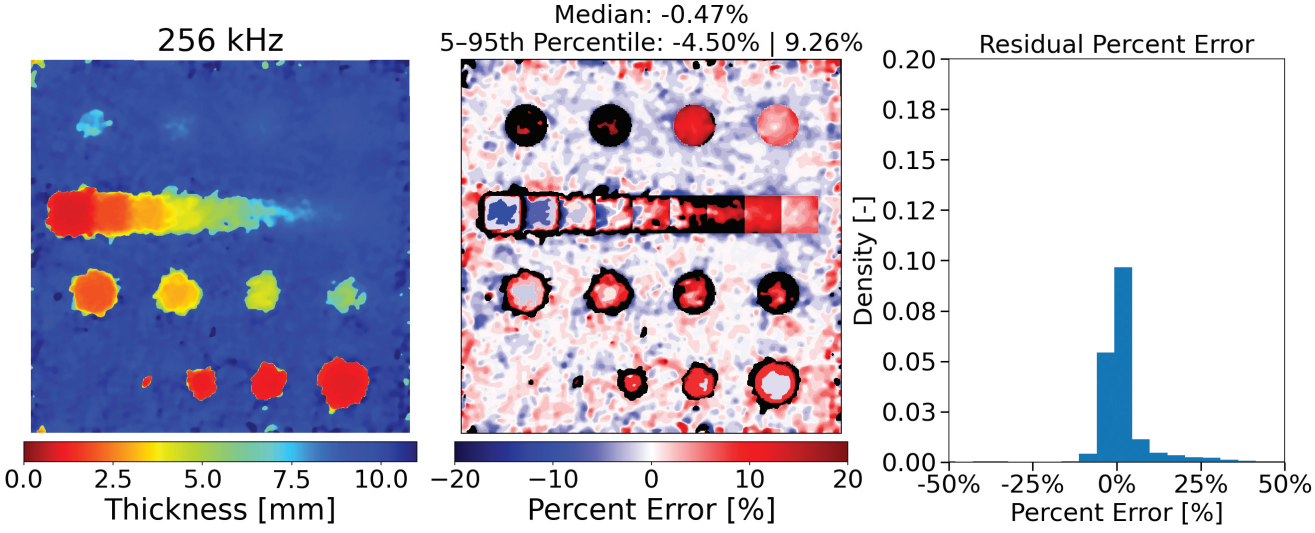


Fig. 6 The most accurate single frequency thickness estimate from the experimental set was collected at 256 kHz.

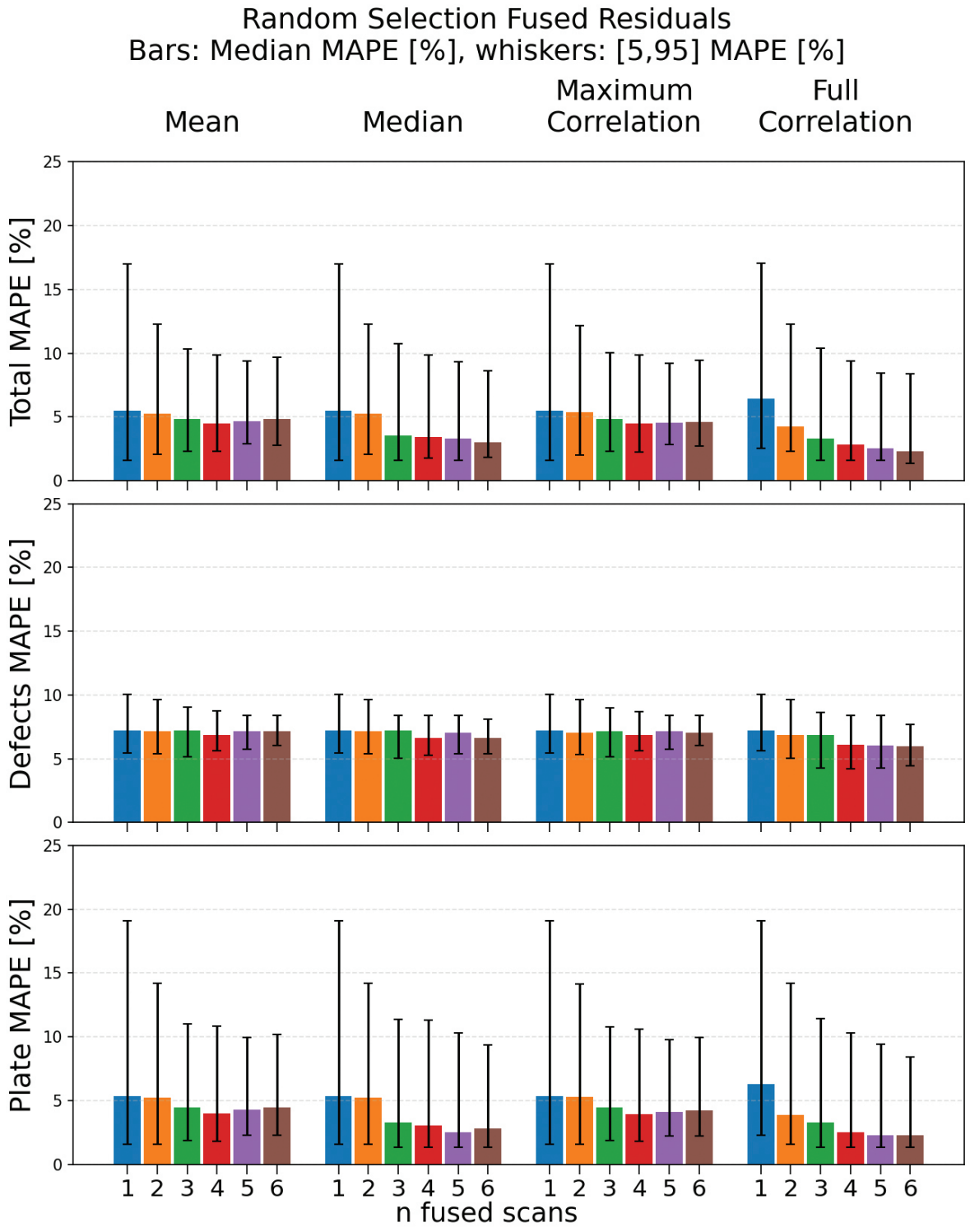


Fig. 7 The distribution of median residuals with increasing the number of scans fused. Each fusion condition is evaluated on the same 100 sets of randomly selected scans.

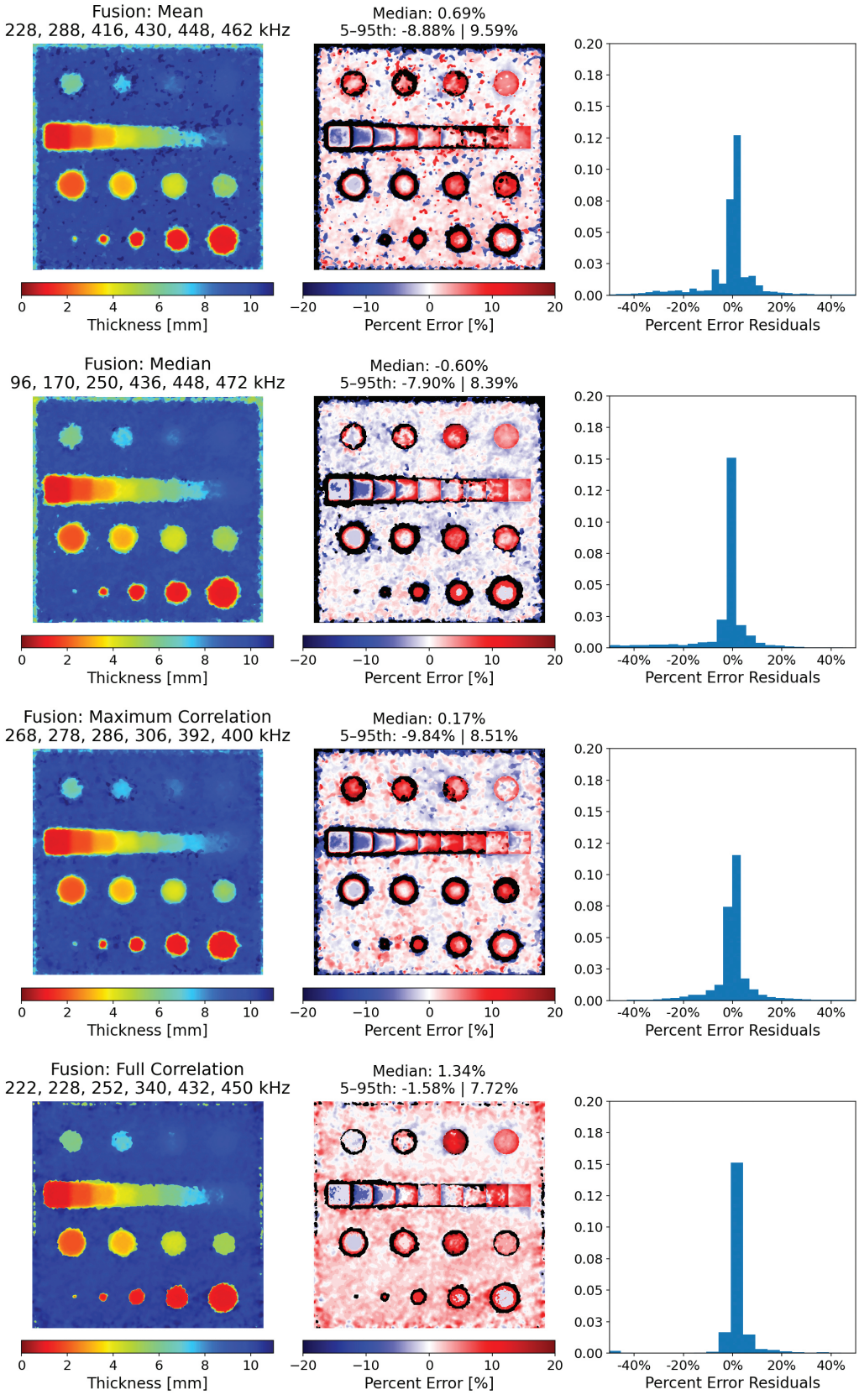


Fig. 8 The distribution of median residuals with increasing the number of scans fused. Each fusion condition is evaluated on the same 100 sets of randomly selected scans.

and narrower defects. As compared to the single scan baseline, percent error residuals had a much tighter distribution while the full correlation displayed the narrowest range, but resulted in slight overestimation. Another interesting point remains in the selection of scans for good performance. The median mechanism was the only method to use scans with values below 200kHz, which were found to be of generally lower quality and resolution. This was likely due to the mechanism of the median, which is effective at disregarding outliers.

Conclusion

This work investigated 4 methods of mearurement fusion for improving thickness estimation on an experimental dataset containing steady-state wavefield measurements that ranged from 50 kHz to 500 kHz in excitation frequency. Using the full correlation curve when combining estimates resulted in the best performance for resolving the nominal plate and full range of thickness reductions in the test sample. Because this method was tested on a dataset with varying scan quality, this performance indicated a higher robustness as compared to the other methods, which is attributed to its incorporation of more correlation information beyond just the maximum value. These results indicate a promising direction for maximizing the potential of analytic approaches to the estimation of thickness from guided-wave analysis.

References

1. Jacobson, E.M., Cummings, I.T., Fickenwirth, P.H., Flynn, E.B., and Wachtor, A.J. "Using steady-state ultrasonic direct-part measurements for defect detection in additively manufactured metal parts". In Epp, D.S., editor, *Special Topics in Structural Dynamics & Experimental Techniques, Volume 5*, pages 59–73. Springer International Publishing (2022).
2. Cummings, I.T., Jacobson, E.M., Fickenwirth, P.H., Flynn, E.B., and Wachtor, A.J. "In-process defect detection for additively manufactured metal lattices". In *Proceedings of the ASME 2020 International Mechanical Engineering Congress & Exposition*, volume 2A, pages 1–8 (2020).
3. Sun, B., Zhang, C., Liu, S., Xu, Z., and Li, L. "Ultrasonic inspection of pouch-type lithium-ion batteries: a review". *Nondestructive Testing and Evaluation*, 39(6):1345–1378 (2024).
4. Yang, Z., Yang, H., Tian, T., Deng, D., Hu, M., Ma, J., Gao, D., Zhang, J., Ma, S., Yang, L., Xu, H., and Wu, Z. "A review on guided-ultrasonic-wave-based structural health monitoring: From fundamental theory to machine learning techniques". *Ultrasonics*, 133:107014 (2023).
5. Flynn, E.B., Chong, S.Y., Jarmer, G.J., and Lee, J.R. "Structural imaging through local wavenumber estimation of guided waves". *NDT & E International*, 59:1–10 (2013).
6. Cummings, I.T., Reinisch, E.C., Jacobson, E.M., Fraser, D.H., Wachtor, A.J., and Flynn, E.B. "Using lidar to identify planar measurement regions in ultrasonic inspections of complex structures". In *Structural Health Monitoring 2021* (2022).
7. Shah, N., Jacobson, E.M., Flynn, E.B., and Wachtor, A.J. "Uncertainty quantification of thickness estimation via full-field ultrasonic inspection". *Structural Health Monitoring 2023*, (0) (2023).
8. Rose, J.L. *Waves in Plates*, chapter 6, pages 76–106. Cambridge University Press (2014).
9. Ostiguy, P.C., Quaegebeur, N., and Masson, P. "Non-destructive evaluation of coating thickness using guided waves". *NDT & E International*, 76:17–25 (2015).
10. Juarez, P.D. and Leckey, C.A.C. "Multi-frequency local wavenumber analysis and ply correlation of delamination damage". *Ultrasonics*, 62:56–65 (2015).
11. Flynn, E.B. "Embedded multi-tone ultrasonic excitation and continuous-scanning laser Doppler vibrometry for rapid and remote imaging of structural defects". In *7th European Workshop on Structural Health Monitoring* (2014).
12. Gao, T., Sun, H., Hong, Y., and Qing, X. "Hidden corrosion detection using laser ultrasonic guided waves with multi-frequency local wavenumber estimation". *Ultrasonics*, 108:106182 (2020).
13. Moon, S., Kang, T., Han, S.W., Jeon, J.Y., and Park, G. "Optimization of excitation frequency and guided wave mode in acoustic wavenumber spectroscopy for shallow wall-thinning defect detection". *J Mech Sci Technol*, 32(11):5213–5221 (2018).
14. Kang, T., Han, S.J., Moon, S., Han, S., Jeon, J.Y., and Park, G. "Lamb-wave sparse-frequency interdigital-transducer-based scanning laser Doppler vibrometry for quantitative depth-wise visualization of defects in plates". *NDT & E International*, 107:102137 (2019).
15. Purcell, F.A.F., Eaton, M., Pearson, M., and Pullin, R. "Non-destructive evaluation of isotropic plate structures by means of mode filtering in the frequency-wavenumber domain". *Mechanical Systems and Signal Processing*, 142:106801 (2020).
16. Eckels, J.D., Jacobson, E.M., Cummings, I.T., Fernandez, I.F., Ho, K., Dervilis, N., Flynn, E.B., and Wachtor, A.J. "Predicting local material thickness from steady-state ultrasonic wavefield measurements using a convolutional neural network". *Ultrasonics*, 123:106661 (2022).
17. Yun, H., Feng, K., Rayhana, R., Pant, S., Genest, M., and Liu, Z. "A multidimensional data fusion neural network for damage localization using ultrasonic guided wave". *IEEE Transactions on Instrumentation and Measurement*, 72:1–12 (2023).
18. Bridge, C.P. "Introduction to the monogenic signal". *CoRR*, abs/1703.09199 (2017).

Chapter 15
Towards Differential Laser Doppler Vibrometer Measurements using Flyable Mirrors for the Vibration Assessment of Remote Structures

Benjamin Halkon and Christian Rembe

Abstract This paper will present findings from the use of Differential laser Doppler vibrometer (D-LDV) measurements from remote targets with the use of UAV-mounted "flyable mirrors". A series of measurements – from the laboratory benchtop through to making real measurements of airborne laser beam steering mirrors – have been conducted to demonstrate, for the first time, the inherent insensitivity of D-LDV measurements to self-noise vibration, either due to the vibration of the sensor head itself or, as is the case here, from vibrating laser beam(s) directing optics. Results are compared in the context of those from single-beam LDV measurements with inertial sensor based "vibration compensation", based on earlier work in which the sensitivity of the measurement to the steering mirror vibration was defined theoretically, using a completely general, vector-based mathematical approach, and rigorously experimentally validated. Target vibration scenarios include those which are statistically stationary and those which are transient, time-varying where frequency and time domain approaches to the correction are utilised. While both approaches are shown to be effective for simple, benchtop scenarios, the D-LDV is shown to offer benefits for more complex, real-world campaigns. However, further work is required to better rigorously and systematically determine the full performance envelope and unique challenges which must be solved before high accuracy, low uncertainty measurement outcomes are realised.

Keywords LDV · UAV · vibration · measurement · infrastructure

Introduction

Laser Doppler vibrometers (LDVs) have been rigorously shown to be as sensitive to vibration of the sensor head as to that of the target itself [1, 2]. For the vibrations observed in common-place measurement campaigns, this sensitivity to this so-called "base motion" can be comparable to that of the target vibration, potentially leading to misinterpretation. For increasingly explored *field-based* vibration measurement capabilities – such as measurements from uncrewed aerial vehicles (UAVs) [3–6] – this particular technical hurdle remains one to be surmounted [7]. Solutions – applicable in both the frequency *and* time domains [8] – have been studied in detail for the purpose of correcting such erroneous LDV measurements using sensor-head mounted accelerometers with excellent results observed[1] [1–3,8]. Nevertheless, accelerometer-based corrections cannot achieve the noise floor and dynamic range characteristics for which LDVs are revered. Practical limitations associated with mounting an LDV sensor-head to a UAV (or "drone") – and recovering useful results from it – persist [7].

Benjamin Halkon
Centre for Audio, Acoustics and Vibration, University of Technology Sydney, 15 Broadway, Ultimo, NSW 2007, Australia
e-mail: benjamin.halkon@uts.edu.au

Christian Rembe
Institute for Electrical Information Technology, Clausthal University of Technology, Adolph-Roemer-Straße 2A, Clausthal-Zellerfeld, D-38678, Lower Saxony, Germany
e-mail: rembe@iei-tu.clausthal.de

[1] with appropriate amplitude and time-alignment approaches implemented for treatment of the correction accelerometer signals

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

As shown schematically and generically in Figure 1, a ground-based, dual, parallel beam *Differential* LDV (D-LDV) has, however, recently been proposed [5] as an alternative to the corrected, single-beam LDV equivalent for the purpose of measuring the vibration of large, remote structures. Here, a drone-mounted "flyable mirror" is used to direct the beam(s) – the path for which is shown with a single, yellow line – from the instrument to the target and back. Meanwhile, a simple but effective, image-based, pan-tilt tracking system can be readily realised to track the mirror from the LDV. The laser beam incidence location *and* direction on the target – both important to control, since the LDV measures the target vibration at the point of incidence in the direction of the laser beam – are clearly a function of both the position and orientation of the mirror *and* the LDV optical axis. Insensitivity to (translational) vibration of the flyable mirror and to spurious additional Doppler signal due to path length fluctuations[2] are the specifically inherent D-LDV benefits in such application scenarios [5, 6].

The former property is the principal subject of investigation herein, where the instrument sensitivity to the laser beam interactions at each of the reflections must be determine. Subsequently, a series of investigations are presented to verify the D-LDV performance and compare it with that of a corrected, single-beam LDV alternative. Ultimately, with these aspects defined, the effective application of a D-LDV to measurements of remote structures using flyable mirrors can be established.

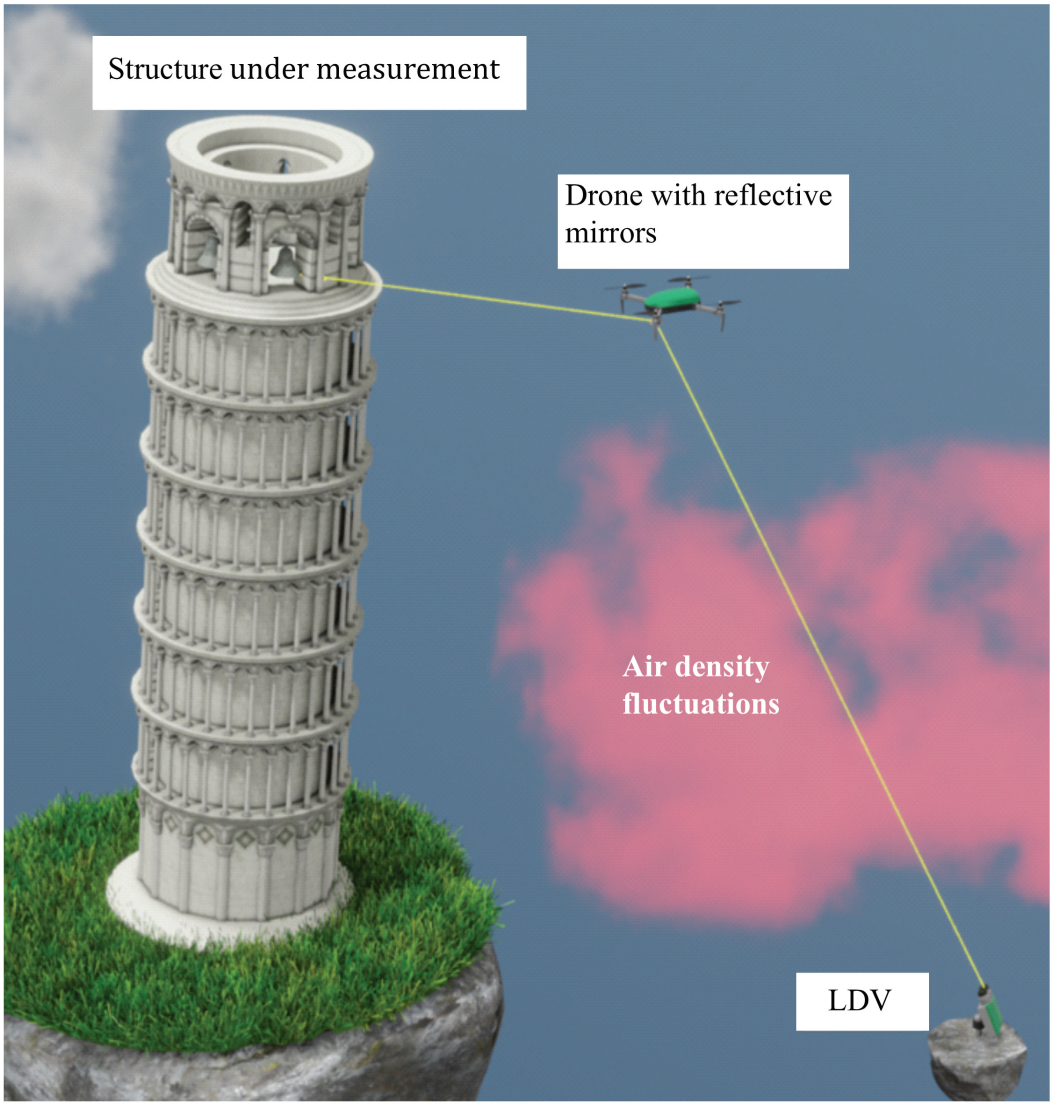


Fig. 1 Schematic LDV measurements from large structures with inaccessible measurement locations of interest [7]

[2] caused by fluid media refractive index between the instrument and the target caused by fluid media refractive index between the instrument and the target

Mathematical Modelling of the Measurement Sensitivity

The total measured velocity in scenarios where vibrating optics are used to steer the laser beam to and from the target of interest is the sum of each of the velocities measured from each vibrating surface at which a change in beam direction occurs. In general, the fundamental relationship between the measured velocity, U_m, and the surface velocity at the point of laser beam incidence, $\overrightarrow{V_P}$, is given by:

$$U_m = \frac{1}{2}\left(\widehat{b}_{n+1} - \widehat{b}_n\right)\vec{V}_{P'} \tag{1}$$

where $\widehat{b}_{n+1}$ and $\widehat{b}_n$ define the directions of the laser beam immediately after and immediately before the laser beam incidence point P', respectively.

For the D-LDV, Figure 2 shows a generalised case where the outbound laser beams are both reflected by a vibrating mirror on their path between the instrument sensor head and the vibrating target. The subscripts $_M$ and $_R$ are used to identify laser beam directions and incidence locations for the Measurement and Reference beams, respectively. In the arrangement shown, the M and R beams are parallel, but this does not have to be the case in general, hence retaining discrete notation for each of the laser beam directions. In alignment with earlier published work [2], further considerations that enable comprehensive modelling of the velocity sensed by each beam are:

a) The points A along the beam path are arbitrarily selected, as is the point B on the mirror surface, which moves to B* following the mirror vibration;

b) While, for simplicity, only one translational and one rotational mirror vibration component are shown, the use of this vector description is completely general, allowing for up to all six degrees-of-freedom of mirror and target vibration;

c) The paths to the target in the absence of mirror vibration (dotted lines) are shown with the beams steered through 90○ which is, again, just an arbitrary choice; and

d) Similarly arbitrarily chosen is the location and orientation of the origin, O, in the plane of the target.

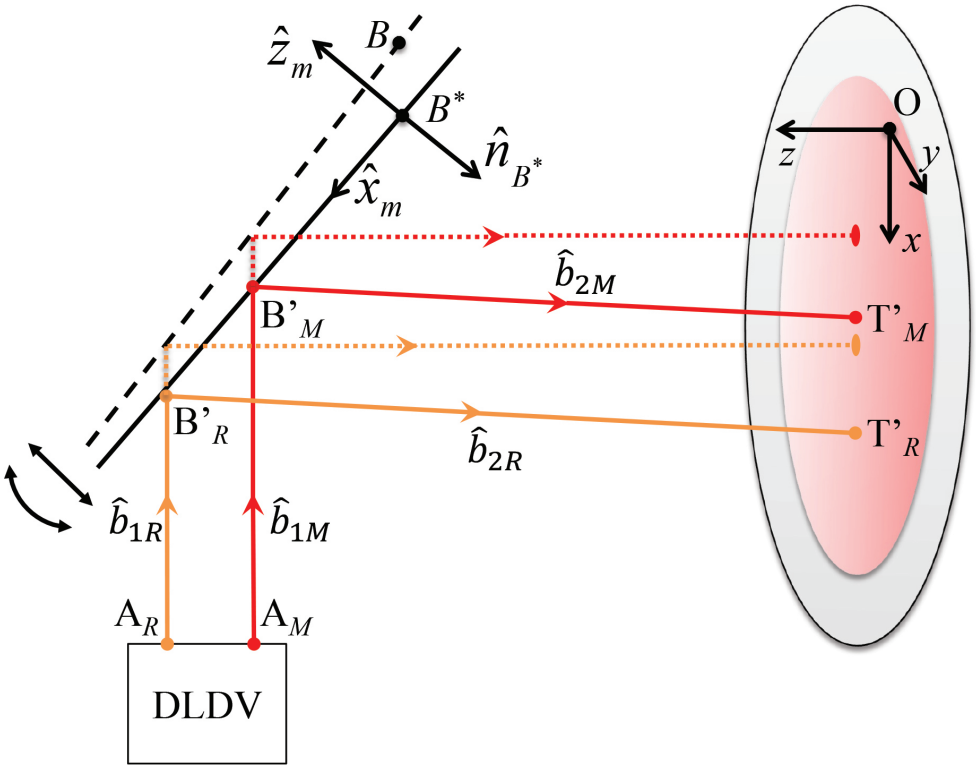


Fig. 2 Schematic showing a generalised D-LDV measurement of target vibration via an angled, vibrating steering mirror [9].

For a conventional parallel beam arrangement - i.e. $\widehat{b}_{1_M} = \widehat{b}_{1_R}$ and $\widehat{b}_{2_M} = \widehat{b}_{2_R}$ - and with the introduction of a non-vibrating retro-reflector in the target plane at the point T′R[3], it can be shown [10] that the measurement is given by:

$$U_m = -\widehat{b}_2 \cdot \overrightarrow{V_{T'_M}} + \left(\widehat{b}_2 - \widehat{b}_1\right) \cdot \left(\overrightarrow{V_{B'_M}} - \overrightarrow{V_{B'_R}}\right) \tag{2}$$

where $\overrightarrow{V_{X'_Y}}$ is the velocity of the point $X(B \text{ or } T)$ for the $Y(M \text{ or } R)$ beam. For scenarios where the mirror vibration is translational only, i.e. where $\overrightarrow{V_{B'_R}} = \overrightarrow{V_{B'_M}}$, this conveniently reduces to:

$$U_m = -\widehat{b}_{2_M} \cdot \overrightarrow{V_{T'_M}} \tag{3}$$

i.e. the intended measurement – the vibration of the target surface at the *Measurement* laser beam incidence in the direction of the incidence beam. This is a very important result since it shows, for the first time in this mathematically rigorous manner without any loss of generality, that the Differential LDV - when configured with a non-vibrating retro-reflector at the target for the *Reference* beam - is inherently insensitive to (translational) mirror vibration.

It should not be a big surprise that such a dual, parallel beam device is insensitive to translational vibration. This is, after all, at the core of the operating principle of the Laser Torsional Vibrometer, first proposed several decades ago [12] and now commercially available and widely deployed for *angular* vibration measurement, especially those associated with rotating shafts and systems. It should be noted that this stationary retro-reflector can be located *anywhere* after the mirror. It is, however, deliberately shown adjacent to the target since the M and R beams takes (almost) the same return path to the target and this delivers the additional benefit that any fluid medium causing path length fluctuations that occur anywhere between the D-LDV and the target are incurred by *both* and are also therefore opto-electronically cancelled [9].

Experimental Investigations

To verify the insensitivity of the D-LDV to steering mirror vibration under controlled conditions, a lab-based, benchtop experimental arrangement as shown in Figure 3 was arranged. The D-LDV and a "check" VibroFlex QTec single beam device are positioned out-of-shot to the lower left of the image field-of-view. A large plane mirror, at nominally 45○ to the laser beams, deflects both M and R beams toward the target plane. The mirror is attached to a large electrodynamic shaker – positioned just out-of-shot to the lower right – with the shaker axis and mirror normal colinear. Vibration of the mirror in the normal direction is realised It is unimportant that the vibrating target – on which the M beam is incident – is a loudspeaker. The signal used to produce the vibration is similarly unimportant, provided that it is analogous in terms of frequencies and amplitude levels to that of a real-world structure. A stationary retro-reflector – on which the R beam is incident – is positioned alongside the target loudspeaker. Reasonable care was taken to align the laser beams, both those from the D-LDV and that from the check LDV, which is directed towards the target via a *stationary* 45○ mirror and approximately incident at/in the same location/direction.

Figure 4 shows an example dataset collected for each of the two LDVs with the shaker-generated mirror vibration also shown for reference. The check LDV captures the true target vibration. While a similar *single-beam* LDV making a measurement of the loudspeaker via the vibrating mirror would be perturbed by the mirror vibration, the D-LDV conversely exhibits no such sensitivity, as can be seen by its very close agreement with the check LDV measurement. The red (D-LDV) curve is almost completely aligned with – and therefore obscured by – the green (check LDV) curve. This is an excellent, expected outcome.

Replacing the mirror/shaker with a mirror attached to a tethered UAV – as shown in Figure 5 – represents an experimental setup with a further level of complication. Here, mirror vibration will certainly be time-varying in nature with a broadband frequency content with peaks associated with the resonances of the vehicle structural and with that of the (inevitably unbalanced) rotors.

Similarly to those data shown and described previously in Figure 4, an example dataset for this experimental scenario is presented in Figure 6. Here, the compact between the D-LDV and the check LDV – in this case targeted directly onto the target loudspeaker from the tripod-mounted instrument (out-of-shot to the bottom right) – is not so evident. While level fluctuations can be seen to occur at similar timeframes, especially from -2.4 to -2.3 s, and underlying signal structures exhibit similar characteristics, differences can surely be attributed to some spurious contribution to the D-LDV measurement as a result of inevitable mirror *rotational* vibration, to which the instrument is *not* insensitive [9]. Future analysis should explore

[3] which may or may not be readily realised yet, for the time being, is a requirement of the proposed approach

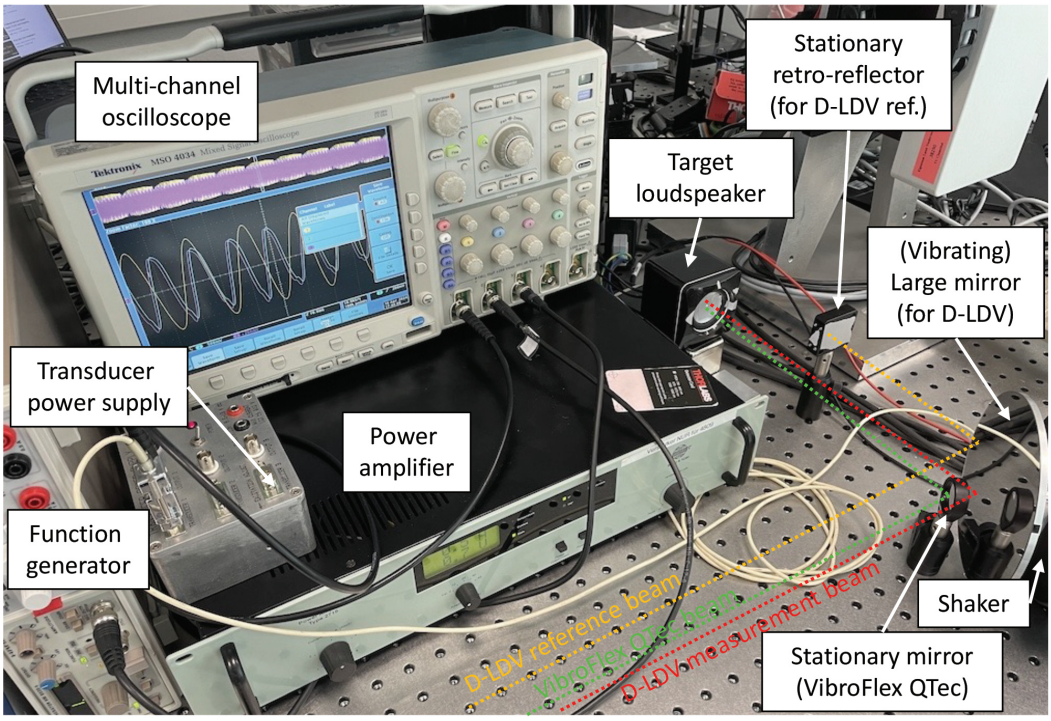


Fig. 3 Experimental setup for a D-LDV measurement from a vibrating target via an instrumented 45○ mirror; laser beam paths for both D-LDV and reference VibroFlex QTec – both out of shot to bottom-left – highlighted for ease of reference [9].

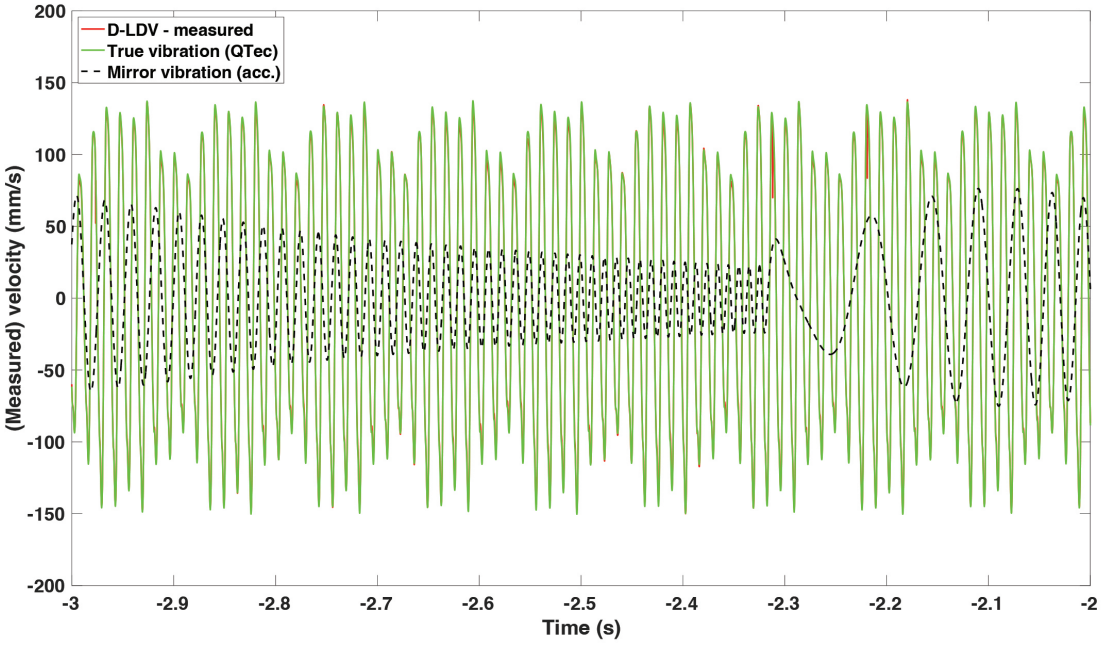


Fig. 4 Example measurement data collected during a D-LDV measurement from a vibrating mirror; 5-100 Hz, 1 s duration chirp – i.e. time-varying – mirror vibration and nominally 75 Hz target vibration [9].

this and the possibility of its correction, for example using an arrangement of transducers on the back-surface of the mirror to extract the required signal.

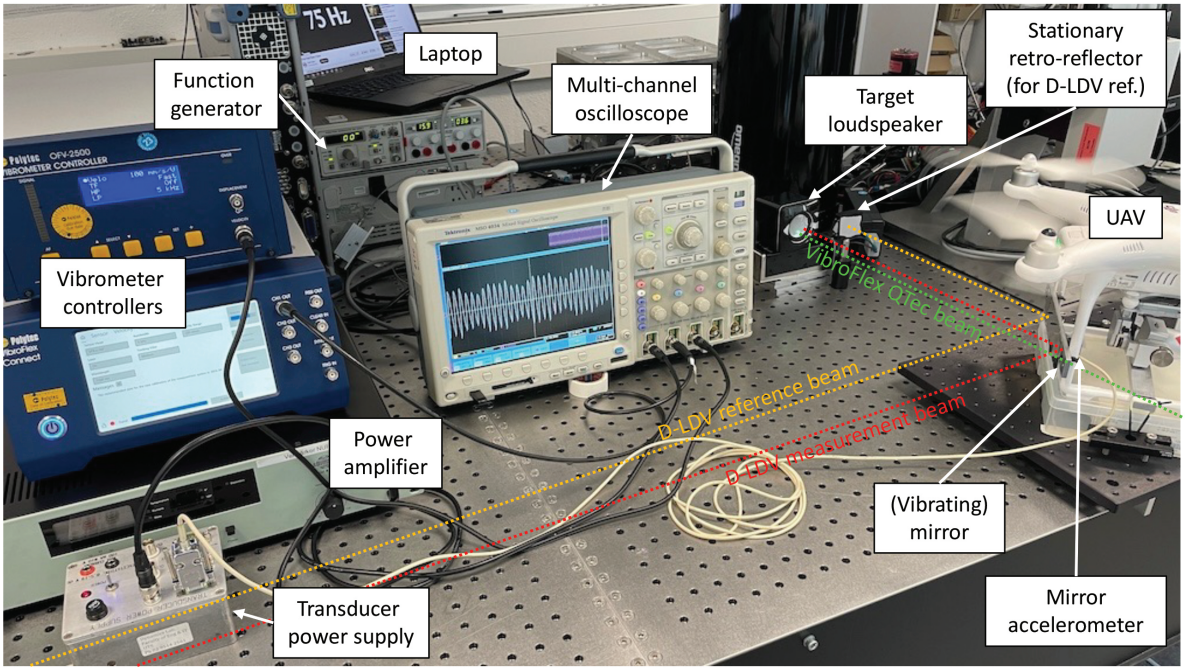


Fig. 5 Experimental setup for a D-LDV measurement from a vibrating target via an instrumented 45○ mirror; laser beam paths for both D-LDV and reference VibroFlex QTec – both out of shot to bottom-left – highlighted for ease of reference [9].

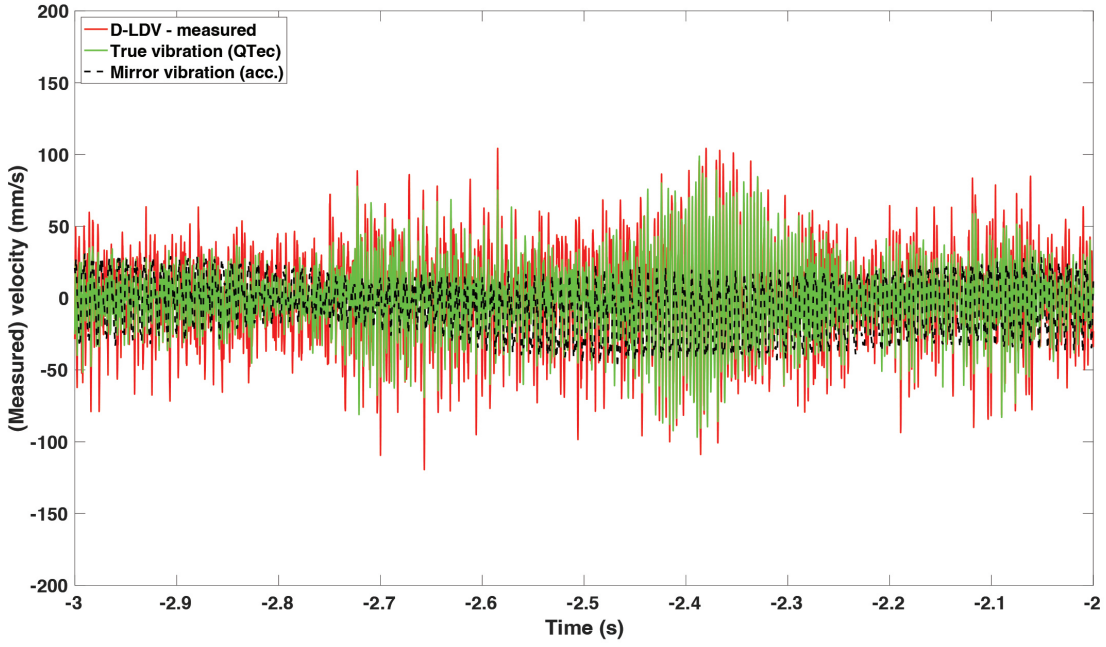


Fig. 6 Example measurement data collected during a D-LDV measurement from a mirror mounted to a tethered UAV; ambient, UAV-generated mirror vibration and transient, time-varying (classical music) target vibration [9].

Lastly, Figure 7 shows a field-based experimental trial where the D-LDV is used to make a remote measurement from a vibrating target via a flyable mirror, as is the ultimate ambition of this emerging capability. While, as before, the target vibration is controlled, here it is realised in the direction of the inbound laser beam using an electrodynamic shaker. The stationary reference beam retro-reflector is adjacent to the target in approximately the same plane. Both surfaces were

relatively large to allow for some beam incidence point variation due to inevitable UAV (whole-body) motion during flight. Such are the practical challenges still to be fully overcome with this, as yet proof-of-concept technique. Figure 8 shows

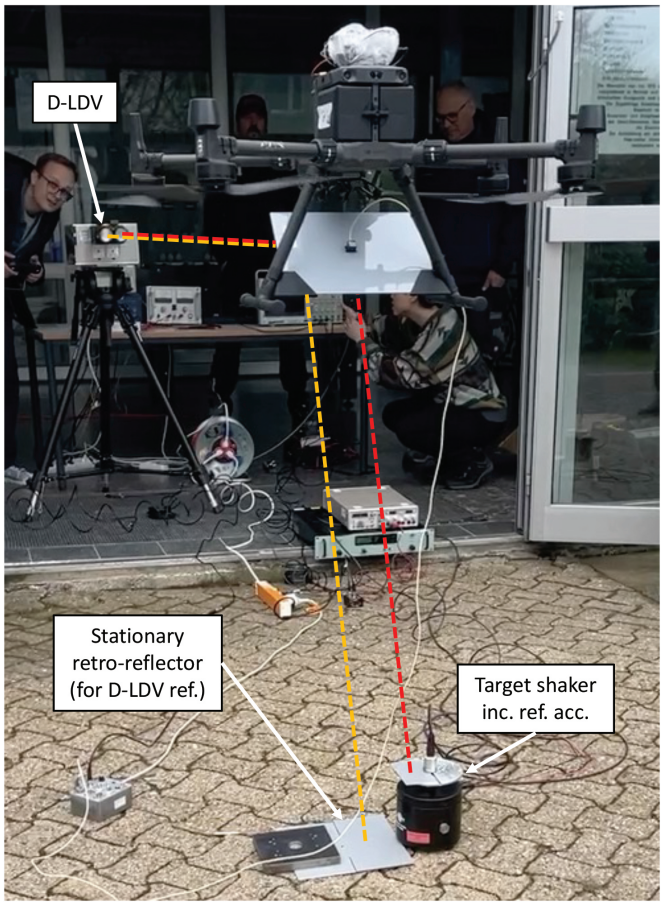


Fig. 7 Experimental setup for a D-LDV measurement from a vibrating target via UAV-mounted flyable mirror [9].

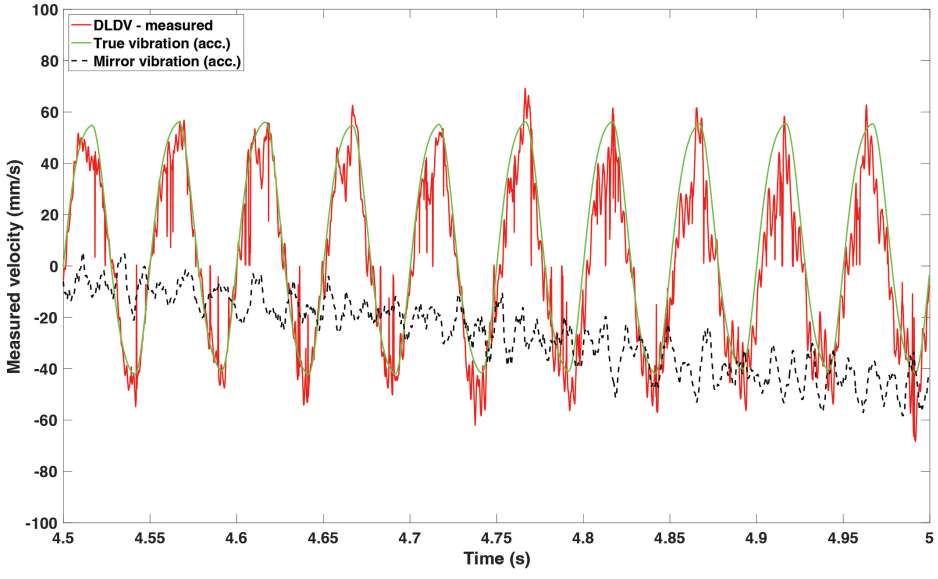


Fig. 8 Example measurement data from an UAV-mounted DLDV for a target vibrating at 20 Hz [9].

an example dataset collected where the 20 Hz (in this case) target shaker vibration is captured quite effectively. Otherwise good agreement between the D-LDV and check measurement – made here with an accelerometer mounted to the target – is impacted by additional signal content with similar characteristics to that of the mirror vibration, surely contributed to by the *rotational* vibration of the mirror.

Conclusions

This paper has presented some of the latest research in the increasingly topical area of LDV measurements of remote structures using UAVs. Rather than directly mounting the LDV to the UAV, combining a ground-based LDV with a "flyable mirror" has the potential to overcome many of the otherwise prohibitive practical challenges. Furthermore, knowing that *conventional* LDVs are as sensitive to sensor head (or beam steering optic) vibration (in the direction of the laser beam) as to that of the target itself, this work focuses on the dual parallel beam *Differential* LDV (D-LDV). Such a device – in which both the Measurement *and* Reference beams take the same path to/from the target of interest – is shown to overcome the self-noise to *translational* vibration as well as offering an inherent means to preclude sensitivity to refractive index fluctuations over the laser beam path length. Increasingly complex experimental set-ups – from lab- to field-based – are used to qualitatively verify a mathematical assessment of the D-LDV sensitivity in the early communication of results offered here. Future investigations will include frequency-domain based representation and quantitative analysis of the D-LDV measured and check signals. Additionally, a means to also mitigate the inevitable sensitivity to mirror *rotational* vibration will be proposed and investigated.

Acknowledgments The authors wish to acknowledge the support of the German Research Foundation (DFG), the German Aerospace Center DLR Institute for Flight Systems – in particular Robyn Zimmermann for UAV-piloting – and the University of Technology Sydney.

References

1. B. J. Halkon, S. J. Rothberg, Reprint of: Taking laser doppler vibrometry off the tripod: correction of measurements affected by instrument vibration, Optics and Lasers in Engineering 99 (2017) 3–10, laser Doppler vibrometry.
2. B. J. Halkon, S. J. Rothberg, Establishing correction solutions for scanning laser doppler vibrometer measurements affected by sensor head vibration, Mechanical Systems and Signal Processing 150 (2021) 107255.
3. B. J. Halkon, S. J. Rothberg, Taking laser doppler vibrometry off the tripod, 25th International Congress on Sound and Vibration 2018, ICSV 2018: Hiroshima Calling (2019).
4. P. Garg, R. Nasimi, A. Ozdagli, S. Zhang, D. D. L. Mascarenas, M. Reda Taha, F. Moreu, Measuring transverse displacements using unmanned aerial systems laser Doppler vibrometer (UAS-LDV): Development and field validation, Sensors 20 (21) (2020).
5. M. Schewe, M. A. A. Ismail, C. Rembe, Towards airborne laser doppler vibrometry for structural health monitoring of large and curved structures, Insight - Non-Destructive Testing and Condition Monitoring 63 (5) (2021) 280–282.
6. M. Schewe, M. A. A. Ismail, R. Zimmermann, U. Durak, C. Rembe, Flyable mirror: Airborne laser doppler vibrometer for large engineering structures, Journal of Physics: Conference Series 2698 (1) (2024) 012007.
7. C. Rembe, B. J. Halkon, M. A. Ismail, Measuring vibrations in large structures with laser-Doppler vibrometry and unmanned aerial systems: A review and outlook, Advanced Devices & Instrumentation 6 (2025) 0103.
8. A. Darwish, B. Halkon, S. Rothberg, S. Oberst, R. Fitch, A comparison of time and frequency domain-based approaches to laser doppler vibrometer instrument vibration correction, Journal of Sound and Vibration 520 (2022) 116607.
9. M. Schewe, D. Kohlmann, H. Wulfmeier, H. Fritze, C. Rembe, Differential laser doppler vibrometry for displacement measurements down to 1 mhz with 1 nm amplitude resolution in harsh environments, Measurement 210 (2023) 112576.
10. B. J. Halkon, C. Gao, R. Zimmermann, C. Rembe, Differential laser Doppler vibrometer measurements using "flyable mirrors" for the vibration assessment of remote structures, Mechanical System and Signal Processing, under review.
11. S. J. Rothberg, M. Tirabassi, A universal framework for modelling measured velocity in laser vibrometry with applications, Mechanical Systems and Signal Processing 26 (2012) 141–166.
12. N. Halliwell, C. Pickering, P. Eastwood, The laser torsional vibrometer: A new instrument, Journal of Sound and Vibration 93 (4) (1984) 588–592.

Chapter 16
Multi Channel Laser Radar for acousto-seimics

W. Zhou, O. Enriquez, G. Pandraud, E. Margallo, J. L. Rubio, and F. Beekman

Abstract We present a novel application of multichannel laser radar (MLR) for non-contact acousto-seismic measurements in granular geomorphic models, targeting early detection of fault propagation and subsurface deformation. Using Ommatidia's Q2 system, featuring 65 phase-coherent channels, we conducted controlled laboratory experiments on sandbanks, simulating fault-like discontinuities. Broadband excitations were introduced to generate elastic wavefields, with the Q2 system capturing simultaneous amplitude and phase responses across the measurement domain.

This dense spatial sampling enables wavefront inversion, allowing reconstruction of propagation paths, dispersion characteristics, and localized strain anomalies associated with evolving fault structures. The study highlights the capability of MLR to overcome limitations of traditional contact-based sensing by providing high-fidelity, real-time vibration mapping without mechanical perturbation of the medium. Embedded SpeckleGuard^{TM} signal diversity further enhances measurement robustness under challenging granular scattering conditions.

Our results demonstrate that multichannel laser radar can serve as a powerful tool for non-invasive subsurface imaging and real-time deformation monitoring in analogue geomechanical models.

Keywords Multichannel Laser Radar · Acousto-Seismics · Wavefront Inversion · Geomorphology · Fault Detection · Non-Contact Vibration Measurement

Introduction

Ommatidia's Q-series Multichannel Laser Radar (MLR) has already been successfully deployed in civil engineering and structural health monitoring (SHM) applications [1, 2]. However, its extension to high-resolution distributed systems for geomechanical studies is a more recent development. This work investigates whether such an emerging non-contact technology can be effectively applied to acousto-seismic analyses, leveraging the latest 65-channel MLR to enable subsurface fault detection in granular geomorphic structures. Granular materials—such as sand—serve as analogues for studying fault dynamics, as their response to stress or excitation provides valuable insight into subsurface structure evolution.

Driven by the limitations of mechanical contact sensors and the need for real-time, minimally intrusive monitoring in laboratory-scale geomorphological models, our approach integrates broadband excitation with synchronized spatial acquisition. Previous work employing Laser Doppler Vibrometry (LDV) on dry granular materials (e.g., Bodet et al. [3], Zaccherini et al. [4]) used mechanical shakers to generate acoustic waves within the sample. However, inserting a shaker into the medium causes structural deformation during experiments and constrains the mobility of the source, thereby limiting the ability to perform comprehensive spatial surveys.

Building on these efforts, [5] demonstrated a non-contact approach in which an acoustic wave is generated at the surface via a remotely pulsed laser, with signals subsequently recorded by a separate LDV. While such configurations successfully capture surface wave sensitivity to small elasticity changes, they remain sequential and alignment-dependent, increasing

W. Zhou · F. Beekman
Utrecht University, Earth Simulation Lab, Princetonlaan 4, 3584 CB Utrecht, The Netherlands
e-mail: w.zhou@uu.nl; w.w.w.beekman@uu.nl

O. Enriquez · G. Pandraud · E. Margallo · J.L. Rubio
Ommatidia LiDAR, Faraday Street 7, 28049 Madrid, Spain
e-mail: oscar-enriquez@ommatidia-lidar.com; gregory@ommatidia-lidar.com; eduardo@ommatidia-lidar.com; jose@ommatidia-lidar.com

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

experimental complexity. By contrast, our proposed method significantly simplifies acquisition by enabling simultaneous multichannel measurements across the entire field of view—eliminating the need for scanning mechanisms and precise repositioning. In the following sections, we focus on sand, a representative granular medium that presents notable challenges in data acquisition and post-processing due to speckle effects and heterogeneous scattering.

Test Set up

In the current work the experimental setup the Q2 MLR system is placed at 90° above test media, allowing a complete coverage of the sandy zone (Figure 1). Current set ups have been used in measuring Quay walls in Amsterdam at a fixed position [2]. The sand bank is composed of two sections; the separation mimicking a fault. Measurements were conducted across the fault with an excitation placed in the sand and applied to study natural fault mechanics. In some cases, one section of the sand medium was loaded to introduce artificial discontinuities. The vibrational response was captured across 65 acquisition points, providing phase-coherent datasets suitable for inversion.

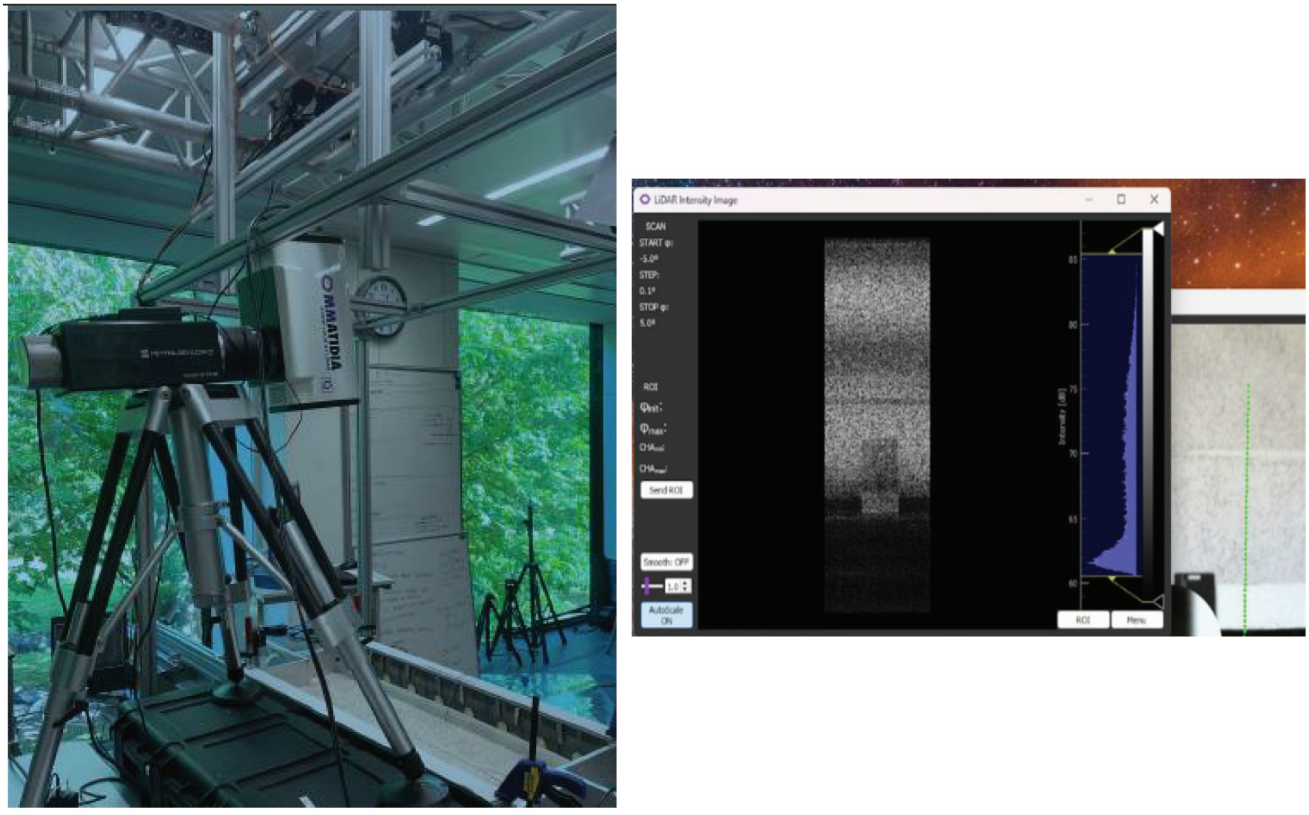


Fig. 1 *Right:* The Q2 Multichannel Laser Radar (MLR) positioned directly above the sandbank during testing, providing full overhead coverage of the experimental domain. *Left:* RGB image of the sandbank surface showing the distribution of the 65 measurement channels across the fault zone.

Measurement Procedure for Granular Material Models

To monitor the deformation of analogue models during an experimental run, we mimic (see Figure 1 left) the concept of seismic imaging widely used in field subsurface exploration at laboratory scale. In field studies, a source creates an elastic wavefield that travels into the subsurface. Reflected, refracted and diffracted waves are subsequently detected by a geophone array placed at the surface. By collecting data for a sufficiently large combination of source and receiver positions, inferences about medium parameters (e.g. *P* and *S* wave speeds) as wells as a reconstruction of the subsurface can be made [6].

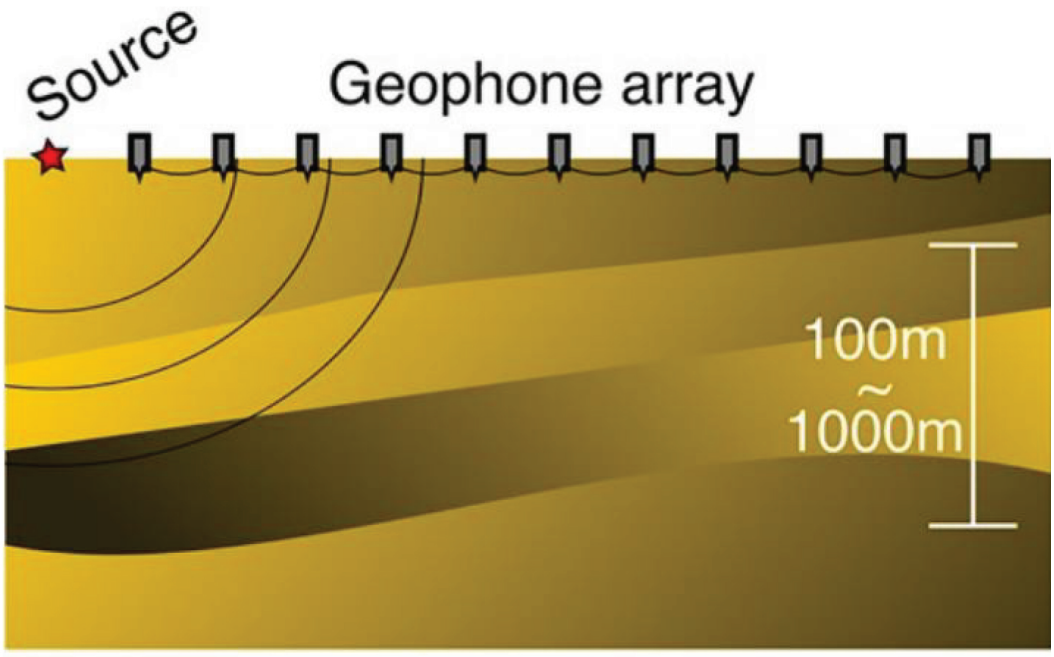


Fig. 2 Acousto-seismic field study

Experimental Results

We report here two types of experimental measurements designed to investigate the influence of fault geometry and deformation on wave propagation within the granular medium. In the first configuration, the sandbank remained stable, allowing us to analyze how the fault structure affects the propagation and reflection of elastic waves under controlled conditions. In the second configuration, a portion of the sandbank was mechanically displaced, altering the geometry of the fault to simulate an earthquake-like event and assess the corresponding changes in wave behavior.

Fixed configuration

In the stable configuration, the Q2 Multichannel Laser Radar (MLR) operated in its active excitation mode, generating a linear chirp signal ranging from 1 to 5 kHz over 1 s with an amplitude of 200 mV. A series of single-line measurements was performed across the fault to validate system accuracy and assess baseline response characteristics. Even without advanced post-processing, these initial scans revealed discernible differences in amplitude and phase across the discontinuity.

Applying wavefront inversion techniques to these datasets—commonly used in geomorphological studies to infer how disturbances propagate through space and time—enabled the mapping of fault propagation paths and the identification of high-strain zones (Figure 3). The reconstructed wavefields show strong agreement with acousto-seismic field analogues, exhibiting consistent arrival times, wave mode separation, and dispersion patterns, thereby validating the capability of the Q2 system for high-fidelity, non-contact subsurface imaging in granular media.

Shear test

In this second experiment, the same excitation parameters as in the stable configuration were maintained—namely, a 1–5 kHz chirp with 1 s duration and 200 mV amplitude. Measurements were conducted along a single line crossing the fault zone over a 10-second acquisition period. During the test, one side of the sandbank was gradually displaced, introducing a shear deformation that altered the fault geometry and mimicked an earthquake-like slip event.

The resulting measurements revealed that the granular surface produced significant speckle-induced noise in the raw data (Figure 4, left), a common challenge in laser-based vibrometry on rough, heterogeneous media such as sand. However, the Q2 system's embedded signal diversity module, SpeckleGuard^{TM}, effectively mitigated these artifacts by combining multiple phase-diverse signals. This enhancement substantially improved measurement accuracy and sensitivity, resulting in clearer and more reliable vibrometric data suitable for subsequent analysis and inversion.

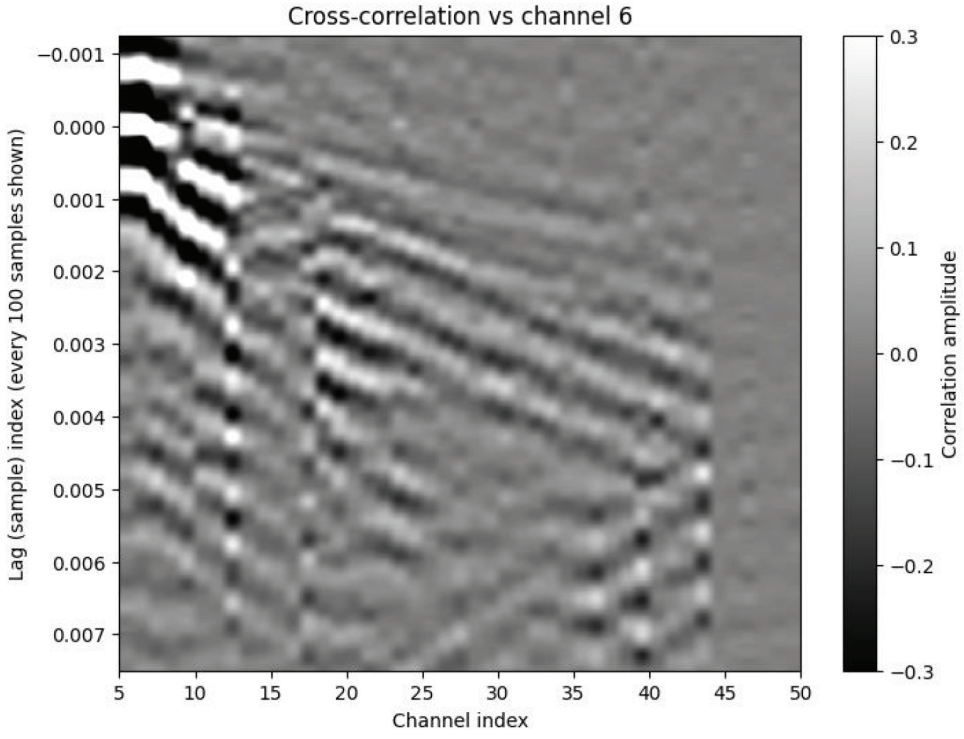


Fig. 3 Fault amplification and reflection at the boundary as extracted from a single line measurement.

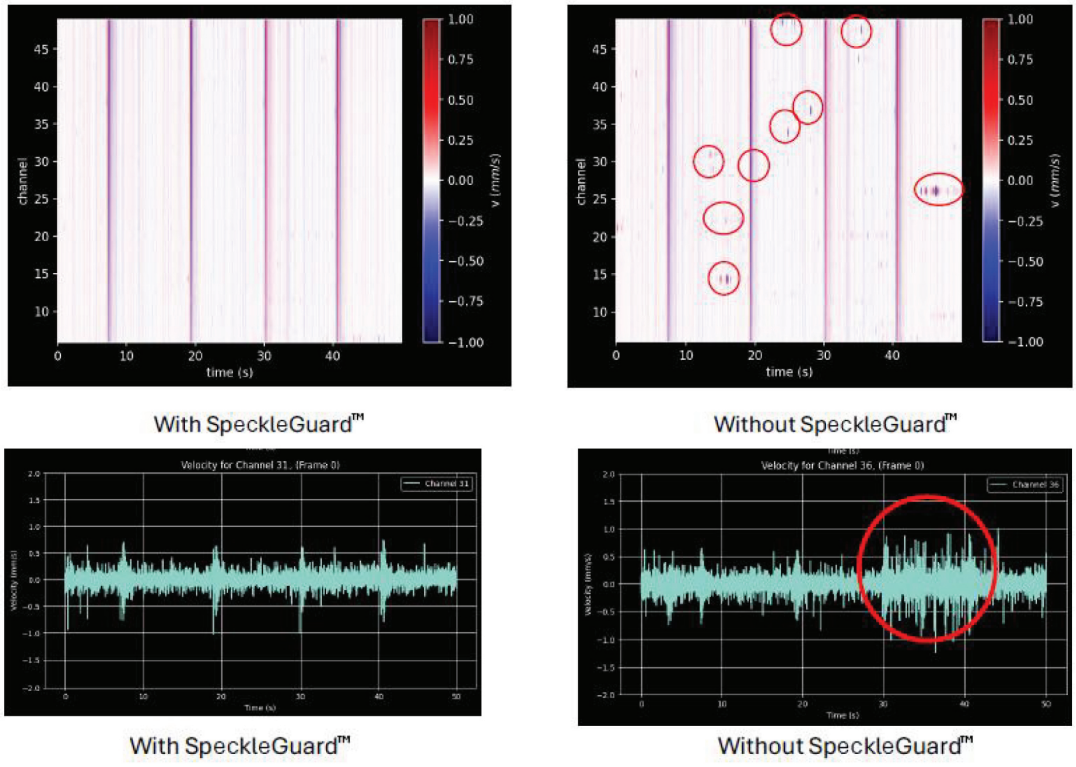


Fig. 4 Effect of signal diversity on vibration measurements acquired over the sand surface. *Top:* Raw signals recorded across all channels, showing the strong speckle-induced noise typical of granular media. *Bottom:* Processed signal from Channel 36, located on the opposite side of the fault relative to the excitation point, after applying the SpeckleGuard^{TM} diversity algorithm, illustrating the improved clarity and signal fidelity achieved through speckle mitigation.

Conclusion

This study confirms the effectiveness of high-resolution, non-contact LDV for structural analysis in granular geomorphic models. The proposed method yields dense vibration datasets suitable for reconstructing subsurface elasticity maps and fault evolution in near real time. Combining sensor innovation with robust signal processing, our platform enhances the ability to monitor and understand deformation processes in both Earth analogue and planetary environments.

References

1. O'Brien, S.A. *et al.* (2025). High-Sensitivity Vibration Monitoring of La Marota Viaduct Using a Novel 128-Channel Laser Radar System. In: Caetano, E., Cunha, Á. (eds) Experimental Vibration Analysis for Civil Engineering Structures. EVACES 2025. Lecture Notes in Civil Engineering, vol 674. Springer, Cham. https://doi.org/10.1007/978-3-031-96110-6_9
2. ter Meulen, D.W.B., Cabboi, A., Brownjohn, J.M.W., Antonini, A. (2025). Vibration-Based Monitoring of the Historic Quay Walls of Amsterdam. In: Cunha, Á., Caetano, E. (eds) Experimental Vibration Analysis for Civil Engineering Structures. EVACES 2025. Lecture Notes in Civil Engineering, vol 676. Springer, Cham. https://doi.org/10.1007/978-3-031-96114-4_9
3. Bodet L., Dhemaied A., Martin R., Mourgues R., Rejiba F., Tournat V., 2014. Small-scale physical modeling of seismic-wave propagation using unconsolidated granular media, Geophysics, 79(6), T323–T339..10.1190/geo2014-0129.1
4. Zaccherini R., Palermo A., Marzani A., Colombi A., Dertimanis V.K., Chatzi E.N., 2022. Geometric and material attenuation of surface acoustic modes in granular media, Geophys. J. Int., 230(1), 288–297..10.1093/gji/ggac076
5. J Smits et al., Non-contacting laser-based acousto-seismics at the laboratory scale: towards near-real-time monitoring of granular analogue models, *Geophysical Journal International*, Volume 238, Issue 1, July 2024, Pages 485–495,
6. Sheriff R.E., Geldart L.P., 1995. Exploration Seismology, 2nd edn, Cambridge Univ. Press.10.1017/CBO9781139168359

Chapter 17
Optimization of Flatbed UV-Printed Digital Image Correlation Speckle Patterns for Microscale Strain Mapping

Aastha S. Singh, Aaron P. Stebner, and Zachary D. Brunson

Abstract Digital image correlation (DIC), an optical method of measuring full-field deformation, relies on the use of applied speckle patterns to track strain, with the quality of the pattern directly influencing the error and variability of the results. Traditional methods of pattern generation lack the resolution and control necessary to pattern small samples, but using UV printers to print computer-generated speckle patterns may offer a solution, seeing that the placement, size, and shape of speckles can be precisely controlled. A computational study was run to investigate how these parameters affect the performance of the pattern under a variety of different tensile, compressive, and shear strain states. Results indicate that irregularly shaped speckles with some randomness provide the most accurate and consistent patterns with which to measure axial strain, but no pattern was able to correctly convey shear strain. Randomness in speckle placement was found to influence precision more than accuracy, seeing that grid-like patterns with no positional randomness had significant variability. UV printing these patterns reveals that imperfections caused by the spreading of printer ink before curing may introduce enough inherent randomness to make positional randomness redundant.

Keywords DIC · Speckle Pattern · Randomness · Repeatability

Introduction & Background

Digital image correlation (DIC) is a non-contact method of measuring strain in which a pattern of dots, called a speckle pattern, is applied to a specimen and digital image registration is used to track the movement and deformation of the dots as the specimen undergoes deformation. By tracking sub-regions of the speckle patterns, called subsets, and comparing them to prior and future pattern images, a full-field strain measurement of the specimen may be obtained with the assumption that the pattern and the specimen are fully bonded and move as one body [1, 2]. DIC is used to examine a multitude of materials, including metals, plastics, composites, and biological materials, under both quasi-static and dynamic loading conditions, steadily replacing other techniques such as various interferometry techniques and other optical methods, such as Moiré patterns [3].

The performance of DIC measurements often rests on factors relating to the speckle pattern itself [4, 5]. The orientation and distribution of the speckles, or the randomness of their placement, is often a key factor in creating a high-quality pattern – uniform grids of patterns often result in false matching, due to correlation software mistaking strain for rigid body motion when tracking the movement of a speckle. However, too much randomness may result in large uniform black or white regions, leading to areas of data loss. Speckle density is also considered, with a 50% coverage being considered ideal to maximize contrast and data collection points [1, 6, 7]. Speckles themselves should be around 3-5 camera pixels across – sizing speckles too small leads to data loss from lack of camera resolution and sizing them too large leads to unnecessary data collection, taking up processing time and storage space. Pixels may also be susceptible to bleeding light into neighboring

Aastha S. Singh · Zachary D. Brunson(✉)
George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, 801 Ferst Dr. Atlanta, GA 30318
e-mail: asingh913@gatech.edu; zachary.brunson@me.gatech.edu

Aaron P. Stebner
George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, 801 Ferst Dr. Atlanta, GA 30318
School of Materials Science and Engineering, Georgia Institute of Technology, 801 Ferst Dr. Atlanta, GA 30318
e-mail: aaron.stebner@gatech.edu

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Computer Vision and SLDV for Structural Dynamics, 2026, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814177

pixels the smaller they get, so subsets that are smaller than necessary may increase error without benefit. The subset size is crucial to the accuracy of the strain measurement as well. Subsets that are too small risk not having enough information to differentiate themselves from neighboring subsets, or even not having enough information to differentiate displacement. Conversely, overly large subsets result in deformation information being averaged over a larger area, making displacement gradients more difficult to discern. Both result in poor strain measurements and high error [1, 6–8].

Traditional DIC pattern techniques include spray paint, rollers, or colored powders applied randomly to specimen surfaces, with the most accurate and viable methods being paint- or ink-based [8, 9]. While traditional DIC patterning techniques are sufficient for larger samples, standard DIC systems are not able to conduct measurements on samples smaller than 10mm without specialized microscopes [10]. For materials that are expensive, delicate, or difficult to manufacture, testing large samples of materials may be costly or extremely difficult, so traditional DIC methods may not be a feasible option for testing, seeing that traditional speckle creation methods lack the precision and resolution necessary for reliable analysis at small scales. Traditional methods struggle to achieve speckle uniformity and often push specimen coverage well beyond the ideal 50% threshold, due to an inability to control speckle size, while novel methods of printing optimized patterns directly onto samples yield much better accuracy and precision [8, 9].

Currently, there is no method to generate highly controlled and repeatable speckle patterns, but utilizing UV printers may be able to address this gap by creating patterns with precise control over parameters like speckle size, density, and placement [9, 11]. In addition to optimizing these parameters, UV printers offer the possibility of creating patterns that are isotropic and scaled to match the gauge size of any given specimen, including ones smaller than 10mm. Research looking into engineered patterns is sparse and UV printing for DIC applications has been seldom explored. In one of the few studies in which UV printed patterns were appraised, the patterns yielded were deemed to be too large to analyze [11, 12].

While prior work has explored the feasibility of using DIC to measure deformation at micro-scales and the performance of computer-generated DIC patterns, no study has examined the merits of computer-generated UV-printed DIC speckle patterns [11, 12]. Thus, this study endeavors to (1) evaluate whether UV-printed patterns can accurately track deformation in small specimen while (2) examining the tradeoff between controlling patterning and randomness to improve data density, improve measurement accuracy, and reduce error or data loss.

Modeling Methodology

To simulate UV printed speckles prior to physical experiments, a computational study was run to evaluate how different parameters affect the performance of speckle patterns. A MATLAB script was developed to create and deform the speckle patterns under seven different applied strain states.

Pattern dimensions were set to be 4.5 mm $\times$ 1.8 mm, corresponding to the planned gauge size of future experimental follow-up investigations. The patterns were designed to fit the resolution capability of a Roland VersaUV LEF-12i Benchtop UV Flatbed Printer, which has a printing resolution of up to 1440 $\times$ 720 dots-per-inch, corresponding to planned future experiments [13]. These constraints were used to select the speckle density and theoretical camera magnification. In adherence to planned future experiments, patterns were generated as white speckles on a black background.

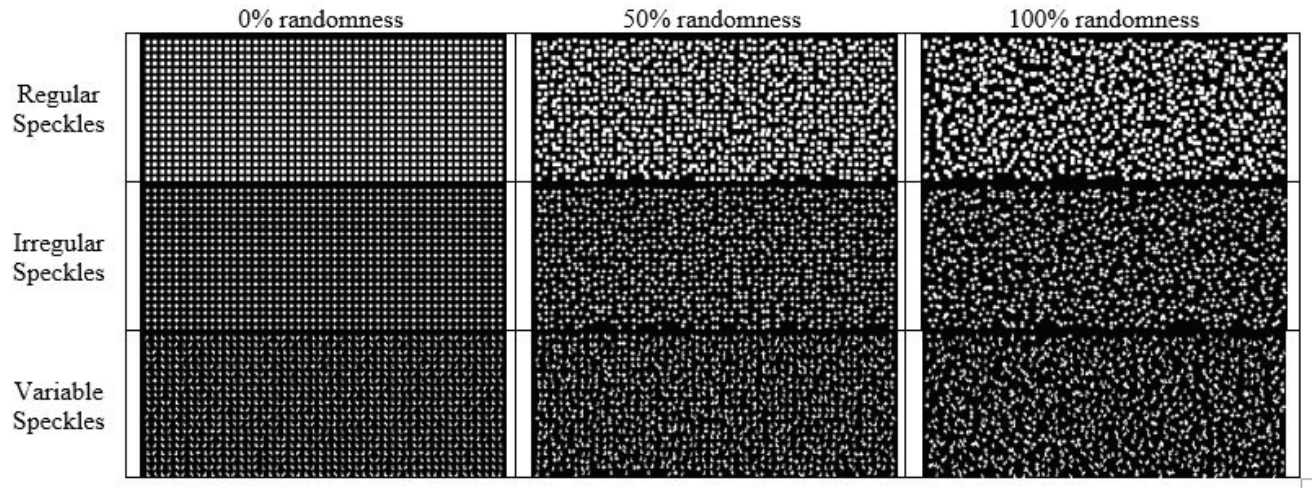
When creating patterns, two different parameters were altered: positional randomness and speckle shape. Positional randomness varied at 0%, 50% and 100%, with 0% randomness being a uniform grid of dots and 100% being a fully randomized spread of speckles. A nominal grid of speckles was created from speckle spacing, size and spacing to decide position. Then, deviation from the grid point was determined based on desired randomness, with more randomness resulting in farther deviations from grid points.

Multiple shape geometries were used to assess the robustness of the patterns. Regular, or box, speckles were the first shape tested; these were uniform square dots with consistent area and boundaries. Irregular, or cross, speckles were also used, which are symmetric shapes with consistent area, but irregular boundaries. The last shape case involved a variety of speckle shapes – crosses, L-shapes, and zig-zag geometries – to represent extreme non-uniformity, with the proportion of each shape in the pattern being chosen randomly using a normal distribution. Speckle density and spacing were decided to ensure that pattern coverage remained as close to 50% as possible to maximize contrast, which was kept uniform for all patterns. All experimental combinations may be seen in Table 1.

Strain was incrementally applied to patterns, based on the input desired number of output images, by identifying coordinates with white ink and multiplying these coordinates by a deformation tensor to get the new location of the speckles, seen as follows [14].

$$\begin{bmatrix} 1+\varepsilon_{xx} & \gamma_{xy} \\ \gamma_{yx} & 1+\varepsilon_{yy} \end{bmatrix} \begin{pmatrix} x_{init} \\ y_{init} \end{pmatrix} = \begin{pmatrix} x_{final} \\ y_{final} \end{pmatrix}$$

Table 1 All speckle pattern parameter combinations tested, in which placement randomness and speckle shape were varied. All other parameters were consistent across tests.



The deformation tensor for each of the seven strain states, along with a brief description, may be found in Table 2. Tensors were divided by the desired number of output images to create steps with which to apply incremental deformation to the pixel locations.

Table 2 All strain state final deformation tensors. Coordinates containing white pixels were multiplied by the tensors to output their new locations under each strain state.

10% tensile strain along x (ε_{tx})	10% compressive strain along x (ε_{cx})	Biaxial tensile strain, 10% along x, 5% along y ($2\varepsilon_{tx} : \varepsilon_{ty}$)	Biaxial tensile strain, 10% along both x and y ($\varepsilon_{tx} : \varepsilon_{ty}$)	10% tensile strain along x and 0.087 radian shear strain along x (ε_{tx} & γ_x)	0.087 radian shear strain along x (γ_x)	0.087 radian shear strain along y (γ_y)
$\begin{bmatrix} 1.1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0.9 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1.1 & 0 \\ 0 & 1.05 \end{bmatrix}$	$\begin{bmatrix} 1.1 & 0 \\ 0 & 1.1 \end{bmatrix}$	$\begin{bmatrix} 1.1 & 0.087 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0.087 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0.087 & 1 \end{bmatrix}$

Testing & Analysis

The computationally deformed patterns were evaluated using VIC-2D 7 to appraise their accuracy and precision under known deformation levels. A comparison of pattern performance across parameters would then offer insight into the combinations that provide the most accurate and least uncertain strain measurements. A subset size of 29 pixels was used with a step size of 7 pixels, as recommended by VIC-2D, to calculate the Lagrangian strain of each test. Figure 1 shows an example of a pattern before deformation, after deformation, and after analysis.

The performance of each pattern under each applied strain state was evaluated using the aggregate statistics of each test analysis, particularly the mean strain values for axial strain in both directions and shear on the XY plane and the standard deviation of these values, a measure of how distant each localized strain calculation was from the global value.

To quantify accuracy, theoretical strain tensors, created from expected strain values for each case, were compared with experimental deformation tensors created using analysis data. For the final frame of each analysis, the theoretical and experimental strain tensors were defined as follows.

$$E_{theory} = \begin{bmatrix} e_{xx,t} & e_{xy,t} \\ e_{yx,t} & e_{yy,t} \end{bmatrix} \qquad E_{exp} = \begin{bmatrix} e_{xx,e} & e_{xy,e} \\ e_{yx,e} & e_{yy,e} \end{bmatrix}$$

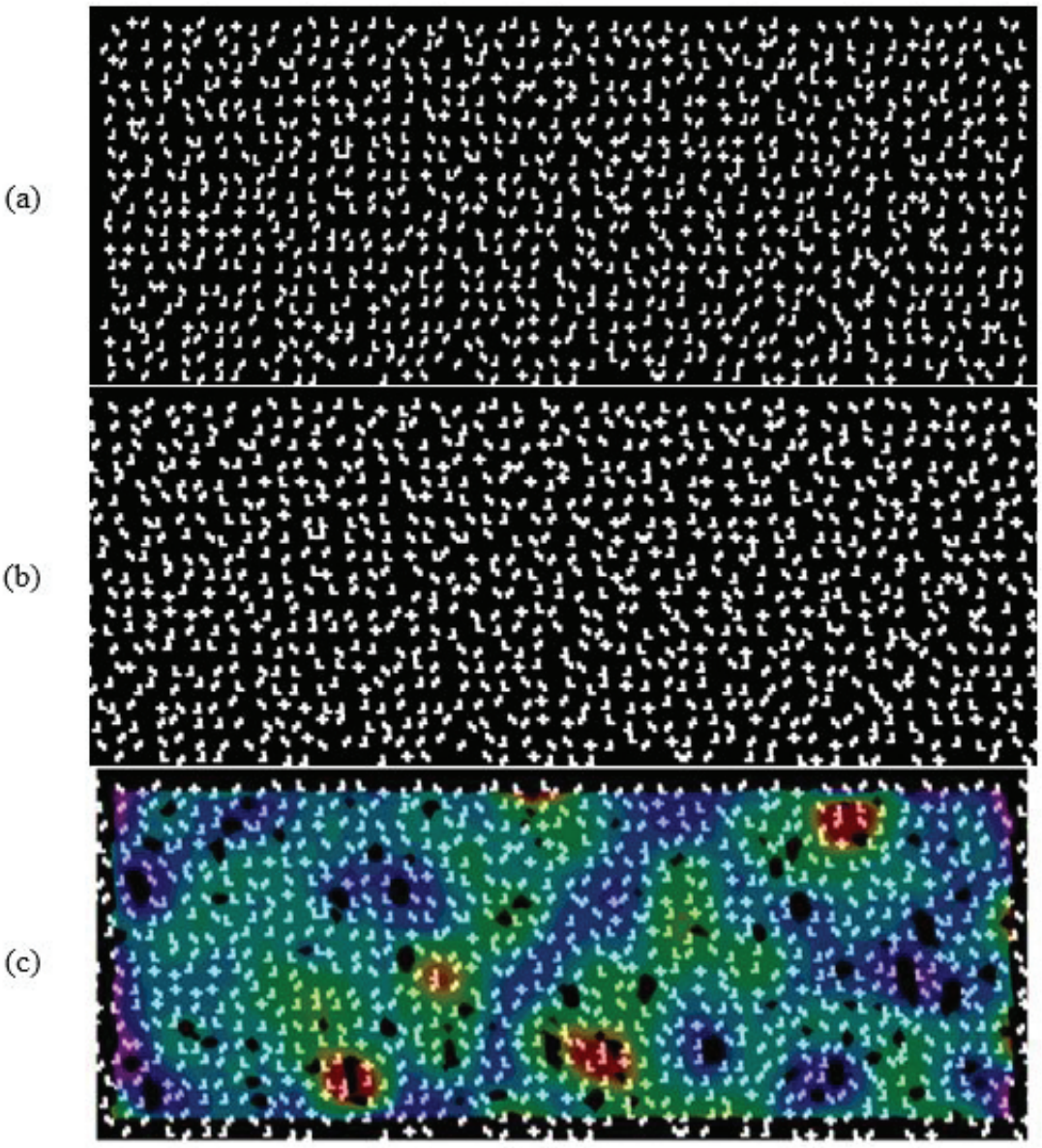


Fig. 1 A speckle pattern with variable speckles and 50% randomness after 10% tensile strain along x with 0.087 radian shear strain along x deformation was applied and then analyzed. (a) shows the image prior to being deformed, (b) shows the image after being deformed, and (c) shows the image after being analyzed.

The error tensor was then defined by the following equation.

$$\Delta E = \begin{bmatrix} e_{xx,t} & e_{xy,t} \\ e_{yx,t} & e_{yy,t} \end{bmatrix} - \begin{bmatrix} e_{xx,e} & e_{xy,e} \\ e_{yx,e} & e_{yy,e} \end{bmatrix} = \begin{bmatrix} E_{xx} & E_{xy} \\ E_{yx} & E_{yy} \end{bmatrix}$$

The final error of each pattern measurement was obtained by taking the Frobenius norm of the final error tensor.

Variability, as a metric of the spread of localized deformation calculations, was calculated by propagating the standard deviations reported.

$$U = \sqrt{(\sigma_{xx})^2 + (\sigma_{yy})^2 + (\sigma_{xy})^2 + (\sigma_{yx})^2}$$

Results

The error of each test as a percentage of the target strain value may be seen in Table 3. Irregular speckles performed the best for axial strain states, with regular speckles also performing well. For regular and irregular speckles, no significant correlation between randomness and performance for axial strain states was noticed except for 1:1 biaxial tension, in which patterns with 0% randomness performed particularly worse. No pattern performed well for any applied shear value. Variable speckles showed acceptable accuracy at most axial strain values for 50-100% randomness, but there was noticeably poor performance across the table for patterns with 0% randomness.

Table 3 Percent error of mean calculated strain relative to theoretical strain for all pattern combinations, varying speckle shape and orientation, and strain states. Color grading highlights performance, as lower values indicate higher accuracy, with green being the most accurate and red being the least.

		ε_{tx}	ε_{cx}	$2\varepsilon_{tx} : \varepsilon_{ty}$	$\varepsilon_{tx} : \varepsilon_{ty}$	ε_{tx} & γ_x	γ_x	γ_y
Regular Speckles	0% randomness	0.76%	1.93%	2.02%	53.42%	14.14%	13.80%	13.87%
	50% randomness	1.26%	1.63%	1.20%	1.64%	15.00%	13.87%	14.09%
	100% randomness	1.37%	1.34%	1.48%	1.89%	14.85%	15.13%	14.54%
Irregular Speckles	0% randomness	0.37%	0.76%	1.28%	14.14%	14.30%	13.57%	13.56%
	50% randomness	1.26%	0.97%	1.48%	1.20%	14.48%	13.71%	13.82%
	100% randomness	1.42%	0.26%	0.69%	1.21%	14.66%	13.72%	13.79%
Variable Speckles	0% randomness	24.78%	13.93%	19.69%	38.32%	18.02%	13.77%	13.84%
	50% randomness	1.72%	13.75%	1.44%	1.68%	14.74%	13.88%	13.87%
	100% randomness	2.15%	13.63%	2.28%	2.73%	14.92%	14.21%	13.79%

The standard deviations of localized strains as percentages of the target strain value are shown in Table 4. Across the board, the irregular speckles showed more stability than either of the other speckle groups - at all randomness levels, strain standard deviation remained below 2.5%. Regular and variable speckles both had cases of standard deviations above 100% and consistently had values above 3%. For these two pattern types, patterns with 0% randomness had noticeably poor consistency, with consistency improving significantly with the introduction of placement randomness.

Table 4 Standard deviations of calculated strain normalized by the target strain value for all pattern combinations, varying speckle shape and orientation, and strain states. Color grading highlights performance, as lower values indicate higher consistency, with green being the most consistent and red being the least.

		ε_{tx}	ε_{cx}	$2\varepsilon_{tx} : \varepsilon_{ty}$	$\varepsilon_{tx} : \varepsilon_{ty}$	ε_{tx} & γ_x	γ_x	γ_y
Regular Speckles	0% randomness	1.56%	2.14%	3.19%	*116.30%*	2.69%	0.90%	0.82%
	50% randomness	1.33%	1.13%	1.55%	2.23%	4.96%	4.69%	3.56%
	100% randomness	1.33%	0.92%	1.76%	1.68%	5.42%	8.34%	4.63%
Irregular Speckles	0% randomness	1.02%	1.49%	1.63%	0.00%	2.38%	0.50%	1.79%
	50% randomness	1.48%	1.18%	1.48%	1.52%	1.37%	0.77%	0.74%
	100% randomness	2.22%	1.88%	2.04%	2.80%	1.84%	1.22%	1.51%
Variable Speckles	0% randomness	*122.28%*	0.52%	*225.54%*	*118.07%*	19.30%	3.08%	0.58%
	50% randomness	1.61%	0.76%	1.47%	1.69%	3.39%	5.95%	1.11%
	100% randomness	2.30%	1.23%	2.50%	2.74%	3.66%	4.53%	1.65%

Patterns were physically printed onto aluminum, the planned experimental test material, spray painted black. Looking at the printed speckles, which may be seen in Figure 2, under a microscope revealed that speckles that were designed to be square digitally exhibited ink bleeding when printed, resulting in distorted edges and higher-than-anticipated white coverage.

Discussion

Speckle shape plays an important role in both the accuracy and precision of DIC pattern performances – dependency is case specific but has the potential to be large. Some cases show that randomness does influence accuracy, but overall, it has a more

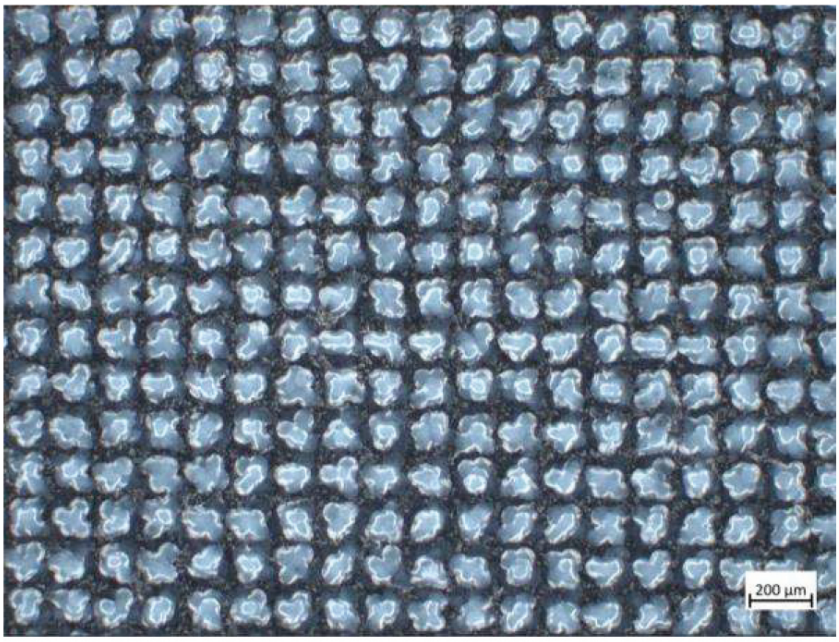


Fig. 2 Printed speckle patterns, intended to be perfect squares, under a microscope to highlight ink spreading and shape irregularities.

significant effect on precision, while the reverse is true for speckle shape. Regular speckles have high accuracy for most axial strain states, except for the 0% randomness pattern and 1:1 biaxial tension strain state. Unexpectedly, randomness had little effect on the error of the strain measurement. Variable speckles introduce significant mean error and error variability in axial strain cases unless adequate randomness is applied, in which case their performance approaches that of irregular speckle patterns. Overall, irregular speckles with non-zero randomness performed the best, providing both accuracy and precision across axial strain states.

Shear strain error was high across all patterns and all stress states, with error values consistently above 13%. This discrepancy indicates that the current computational modeling process may be representative of a DIC setup with too low of a camera resolution to accurately measure shear deformation. One potential improvement, therefore, could be to increase the magnification of the patterns, allowing for better tracking of diagonal displacements.

Overall, the results support the notion that the performance and stability of the strain measurements are roughly correlated with the type of speckles used and the position of the speckles. Irregular speckle patterns with any degree of randomness performed the best overall, showing that irregularity in pattern form improves the robustness of strain measurements. The variability in reported strains worsened across the board for irregular and variable speckle shapes when evaluating 100% random patterns, as compared to 50% random patterns, with no significant impacts on accuracy. Even with regular speckle shapes, no remarkable gains were made in precision when going from 50% to 100% randomness. Patterns with no randomness had remarkably high rates of variability, meaning the strain measurements are unreliable, even with the density of available information.

The bleed of the ink on UV printed test samples suggests that printing patterns may inherently impact the speckle density and randomness of the pattern by introducing randomness in pixel shape, which has the potential to introduce sufficient variability to reach the desired degree of randomness. It remains unknown whether randomness in speckle orientation or shape is necessary – if UV printing the patterns introduces sufficient variability through imperfections caused by ink spreading, the need for deliberate positional randomness may be deemed unnecessary, allowing for higher pixel density and improved analytical resolution.

Conclusion

The results prove that both speckle shape and positional randomness play roles in the optimization of speckle patterns when used in axial strain cases. The accuracy of measurement from a speckle pattern shows a dependence on speckle shape, with

speckles of any standardized shape outperforming variable speckle shapes. The precision of measurement is influenced by the position of the speckles, with positional randomness of any sort decreasing the deformation measurement variability. The degree of randomness, beyond 50%, had little additional benefit, and in fact increased the likelihood of large areas of uniform color, which decreases the amount of information to be derived from that area. Tests with shear strain were inaccurate, indicating the need for higher camera resolution and magnification in future tests.

Printing patterns onto aluminum blanks using a UV printer revealed that the printing process can inherently introduce variability in patterns, which can be utilized to replace positional randomness, possibly improving data density by allowing grids of dots to be used in patterns. Future work will explore the limits and benefits of UV printing for repeatable, customizable, and inherently random speckle patterns at small scales with large camera magnification, allowing the variation in speckles to be seen. Physical validation of selected tensile and shear patterns on aluminum samples will be used to corroborate computational findings, with emphasis during testing and analysis on the effects of magnification, subset size, and step size to improve measurement accuracy

References

1. Sutton, Michael A.; Orteu, Jean-J; Schreier, Hubert W. *Image Correlation for Shape, Motion and Deformation Measurements: Basic Concepts, Theory and Applications.* Springer, 2009.
2. Pan, Bing. "Digital Image Correlation for Surface Deformation Measurement: Historical Developments, Recent Advances and Future Goals." *Measurement Science and Technology* **29**(8), 082001 (2018).
3. Bogusz, Paweł; et al. "Application of Digital Image Correlation for Strain Mapping, Displacement, and Deformation Measurement." *PMC – National Institutes of Health (PMC11174112)*, 2024.
4. Liu, X.-Y.; Li, R.-L.; Zhao, H.-W.; et al. "Quality Assessment of Speckle Patterns for Digital Image Correlation by Shannon Entropy." *Optik* **126**(23), 4206–4211 (2015).
5. Park, J.; Yoon, S.; Kwon, T.-H.; et al. "Assessment of Speckle-Pattern Quality in Digital Image Correlation Based on Gray Intensity and Speckle Morphology." *Optics and Lasers in Engineering* **91**, 62–72 (2017).
6. Correlated Solutions, Inc. *Speckle Pattern Fundamentals.* Application Note, VIC-3D 9, 2021.
7. Dong, Y. L.; Pan, B. "A Review of Speckle Pattern Fabrication and Assessment for Digital Image Correlation." *Experimental Mechanics* **57**(8), 1161–1181 (2017).
8. Pan, B.; Xie, H.; Wang, Z.; et al. "Study on Subset Size Selection in Digital Image Correlation for Speckle Patterns." *Optics Express* **16**(10), 7037 (2008).
9. Zhang, Juan; Sweedy, Ahmed; Gitzhofer, François; Baroud, Gamal. "A Novel Method for Repeatedly Generating Speckle Patterns Used in Digital Image Correlation." *Optics and Lasers in Engineering* **100**, 259–266 (2018).
10. Weissman, Clifford. "Non-Contact 3D Deformation and Strain Measurements Using Digital Image Correlation for Biomedical Applications." *Biomechanics* **2018** (2018).
11. Wang, Y.; Gao, Y.; Liu, Y.; Gao, Z.; Su, Y.; Zhang, Q. "Optimal Aperture and Digital Speckle Optimization in Digital Image Correlation." *Experimental Mechanics* **61**, 677–684 (2021).
12. Kramer, Sharlotte; Reu, Phillip; Bonk, Sarah. *A Speckle Patterning Study for Laboratory-Scale DIC Experiments.* Sandia National Laboratories, Albuquerque, NM, USA, 2017.
13. Roland DG Corporation. *VersaUV LEF-12i Benchtop UV Flatbed Printer: Specifications.* 2020. Retrieved from https://www.rolanddga.com
14. Rudnicki, John W. *Fundamentals of Continuum Mechanics.* John Wiley & Sons, 2014.

River Rapids

Computer Vision and SDLV for Structural Dynamics, Volume 6

Janko Slavič ■ Daniel Rohe ■ Javad Baqersad

Computer Vision and SDLV for Structural Dynamics, Volume 6: Proceedings of the 44th IMAC, A Conference and Exposition on Structural Dynamics, 2026, the sixth volume of ten from the Conference brings together contributions to this important area of research and engineering. The collection presents early findings and case studies on fundamental and applied aspects of Computer Vision and SDLV for Structural Dynamics including papers on:

- Novel Techniques
- Optical Methods
- SLDV Methods
- Photogrammetry & DIC

ISBN 978-87-438-1445-0



River Publishers