Results in Engineering 29 (2026) 108474

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering

Identifying frequency-domain operating deflection shapes and internal damage in structures using geometric vision method

Haoran Zhang [a,1], Jiaren Zuo [a,1], Zhangyu Yao [a], Deyang Zhang [b], Lei Lu [c], Wei Pan [d], Zhilong Su [a,*]

[a] Shanghai Institute of Applied Mathematics and Mechanics, School of Mechanics and Engineering Science, Shanghai Key Laboratory of Mechanics in Energy Engineering, Shanghai University, 200444, Shanghai, China
[b] Ship Scientific Research Center, National Key Laboratory of Ship Structural Safety, 214082, Wuxi, China
[c] Institute for Complexity Science, Henan University of Technology, 450001, Zhengzhou, China
[d] Department of R & D, OPT Machine Vision Tech Co, 523860, Ltd, Dongguan, China

ARTICLE INFO

2000 MSC:
00-01
99-00

Keywords:
Visual deformation measurement
Structural health inspection
Computer vision
Camera calibration
Operating-deflection-shape
Damage identification

ABSTRACT

Operating deflection shape (ODS) analysis characterizes structural vibration in the frequency domain without requiring prior knowledge of excitation, offering an automatic means to assess dynamic behavior and detect damage in service. Building on this, we present a geometric vision–based sensing method for ODS identification under a depth-parameterization framework, enabling direct recovery of full-field vibration responses from spatiotemporal image motion captured by a stereo imaging system. To ensure accurate ODS reconstruction, a random sample consensus–based calibration algorithm is developed to rectify stereo geometry and establish a depth-parameterized projective system. Within this framework, we formulate a linear model that maps measured image motion of object points to their ODS responses in the frequency domain. In addition, a wavelet-based algorithm is proposed to extract damage parameters from defective structures using the reconstructed ODS fields. Experimental validations confirm that the proposed vision-based sensing method and damage identification approach are both effective and robust, providing a promising pathway for structural vibration monitoring and health assessment in complex operational environments.

1. Introduction

Advances in large-scale development of advanced equipment structures in many fields, such as aerospace, marine engineering, clean energy, etc., have been posing increasing demands on dynamic behaviors and damage inspection of structure components during service. Currently, there are many diverse types of sensors built to probe a specific property of structures for fitting such requirements [1]. For example, the vibration sensors that utilize magnetoelastic composites and fiber Bragg gratings (FBG) [2,3]. However, many conventional sensors are bulky when collecting dynamic deformation data of structures in operation due to their contact and monofunctional working mode [1,4]. Even for the well-established non-contact laser vibrometers [5], they are impractical for tasks involving many degrees of freedom (DOFs) due to limitations of single point sampling and the complex system calibration [6].

Over the past few decades, as the demands for non-contact full-field sensing tasks in structural health evaluation and prototyping have increased, and with the advent of advanced optical imaging and processing technologies, two significant trends have emerged in this field to address these challenges. First, it has become essential to introduce computer vision-based measurement techniques that compute displacement or motion in the temporal-spatial domain [7,8]. This trend is evident in the successful application of well-known digital image correlation (DIC), optical flow and target tracking in various fields, including deformation measurement [9–11], vibration measurement using stereo [12,13] or monocular cameras [14–16], modal testing and identification [17–21], and even in prior work on cross-medium static and dynamic 3D shape and deformation sensing [22,23]. Secondly, optical imaging and vision measurement methods have become increasingly important in operating-deflection-shape (ODS) measurement. Many recent advancements have leveraged optical imaging-based ODS inspection algorithms, such as trivial frequency domain triangulation [6] and the binocular structured light measurement method [24].

In contrast to the former category, which reconstructs the motion or displacement fields in 3D Euclidean space with spatial

* Corresponding author.
E-mail address: szloong@shu.edu.cn (Z. Su).
[1] These authors contributed equally to this work

https://doi.org/10.1016/j.rineng.2025.108474
Received 3 October 2025; Received in revised form 22 November 2025; Accepted 27 November 2025
Available online 30 November 2025

triangulation [25], the goal of optical ODS sensing has been proposed to apply two- or multi-view projective imaging to establish a transformation that maps the image motion relationships to the displacement amplitudes in the frequency domain, corresponding to specific vibration frequencies. This approach directly measures the operational vibration response by reconstructing the motion patterns encoded in the high-dimensional image space in the frequency domain. It has an obvious advantage that it is more robust to high-frequency and random noises caused by environmental vibration for allowing us to identify the effective response without losing accuracy. The main reason is the frequency-domain reconstruction can separate the vibration signal, periodic and random noises. In contrast, spatiotemporal domain displacement reconstruction not only includes the displacement amplitude of each vibration mode but also includes noise components, which are often filtered out depending on the local characteristics of the signal while it may lose high-frequency motion information [26]. Besides, analytical reconstruction in the frequency domain gains higher stability and less error accumulation by focusing on the periodic or repetitive features of the dynamic response in the time-domain (such as image motion), providing more precise and detailed response information for downstream tasks, such as delamination defect location [27] and damage identification in this work). Owing to these merits, frequency domain optical ODS sensing has shown a steady and progressive development trend in dynamic deformation characterization.

Recent progress in this field has been driven by innovations in high-speed visual imaging and the implementation of advanced frequency domain triangulation algorithms [6]. We approach these schemes within a uniform framework that is geometric vision ODS sensing. The term "geometric" is intended to emphasize that the dynamic responses of the operating structures can be regarded as points in two respective vector spaces under geometric perspective projection and that the sensing process can be considered as a mapping from one vector space to the other. This mapping is generally constructed based on the well-established Fourier transform. However, the existing methods have certain shortcomings, such as phase delays caused by camera motion and the complexity of time-frequency transformation under perspective projection, highlighting the need for further research to enhance and refine this methodological framework.

The geometric vision ODS sensing framework provides a straightforward solution that allows us to observe the dynamic measures (such as natural frequency and mode shape) in the frequency domain directly. One of the outstanding properties of ODS is that it can characterize vibration responses without needing to know the prior excitation constraints [28]. As a consequence, it also provides an efficient solution to this study for identifying the internal damage of operating structures compared with the existing visual crack detection techniques that target surface defects [29–31]. Traditionally, the damage identification for structures using vibration diagnosis techniques often requires precise excitation information and/or the deployment of contact sensors [32–34]. However, they are affected by the spatial layout of vibration sensors in addition to the excitation and structure priors. Though geometric vision-based method are proposed in the literature more recently for damage localization [35,36], it relies on pseudo-excitation still and/or the calculation procedure is complicated due to the use of some heuristic algorithms such as denoising and clustering. For these reasons, it is essential to develop some effective and practical approaches to resolve the contradiction in probing structural damage in service by combining the strengths of the geometric optical sensing and the excitation irrelevance of ODS.

To this end, we introduce frequency-domain geometric vision ODS sensing, a one-parameter time-frequency transformation method that represents the operating-deflection-shapes with a two-view geometric perspective projection, and then approach damage identification based on the measured ODS data to address the gaps in existing approaches. The contributions of this work are as follows: (1) We formulate the motion relationship from spatial frame to stereo image space as a one-parameter linear mapping under rectified stereo imaging geometry, which is determined through our random-sampling-consistency (RANSAC) calibration; (2) Leveraging this linearity, we propose a novel method that transforms the stereo image motion to the frequency domain directly, enabling instant estimation of full-field ODS with respect to the frequencies; (3) A wavelet-based method is presented to identify the potential damage information embedded in the ODS measures, without reliance on excitation priors. Experiments are conducted to show the performance of the proposed computer vision ODS sensing and the ODS-based damage identification methods. Compared to the existing frequency-domain triangulation methods in [6,37], which are based on single-camera multi-view imaging with relative motion, similar to the structure-from-motion technique (moving camera with fixed object or fixed camera with moving object), our ODS sensing method is established based on stereo-vision geometry, which provides two synchronous views for simultaneously capturing the vibration response without moving the camera or object during the measurement process, thereby eliminating the inherent phase delays caused by the relative motion of a single camera. Moreover, because of stereo vision, our method has a fixed imaging geometry that allows pre-calibrating the imaging system accurately and then applying the one-parameter linear model to directly transform the stereo image motion into ODS in the frequency domain at each step without repeatedly estimating camera poses. This is impossible in single-camera multi-view imaging, where the camera pose varies between views and must be re-estimated for every step of ODS reconstruction, leading to significantly higher computational cost and accumulated calibration errors. Owing to the stereo-vision configuration, our method is also more suitable for engineering applications, as the stereo imaging system requires no manual movement after installation.

The rest of this paper is organized as follows. Section 2 presents the principle of the proposed geometric vision ODS sensing and damage identification framework, including the RANSAC calibration of the two-view stereo imaging system in Section 2.1, the frequency-domain ODS estimation method in Section 2.2, and the damage detection algorithm in Section 2.3. Experimental demonstrations and results are shown in Section 3. We conclude in Section 5.

2. Methodology

In this section, we present the proposed method according to the underlying measurement pipeline as shown in Fig. 1. Consider a stereo imaging system composed of two still high-speed cameras, each of which is amenable to the well-established single-point-of-view projection assumption. We first propose calibrating the imaging system using the RANSAC strategy to mitigate errors caused by imprecise target patterns. Subsequently, we perform stereo rectification on the stereo imaging geometry to establish an efficient ODS estimation model via depth parameterization and Fourier transform. Finally, we show the damage identification algorithm based on wavelet transformation. Details are presented in the following sections.

2.1. RANSAC-Based calibration of the imaging system

For the tasks of high-frequency ODS sensing, the amplitude response is relatively small and is more sensitive to systematic errors caused by calibration of the optical imaging system. Therefore, it is essential to accurately calibrate the camera system before imaging structural vibrations. Though plenty of methods, including pre-calibration [38,39], auto-calibration [40] and the influences [41,42], have been proposed to support system calibration, an obvious problem with current calibration techniques is that accuracy consistency is not easy to guarantee, even for different calibrations in the same environment. For the most widely used Zhang's method [38], which estimates the calibration parameters based on a planar target, the problem arises from the observation that

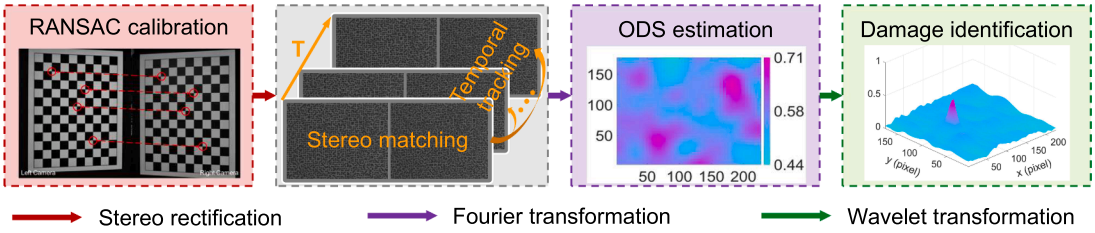

Fig. 1. Method workflow. An overall work pipeline of our geometric optical ODS sensing and damage identification under stereo vision imaging configuration. The pipeline consists of four steps: (1) RANSAC-based calibration of the stereo-vision system, (2) stereo and temporal matching, (3) ODS estimation, and (4) damage identification.

some corners of the planar target may randomly lose imaging precision in different calibration poses, leading to uncertain impacts on the determined imaging parameters such as focal lengths, principal points, and distortion factors. Though specialized imaging or using some special patterns such as deltile grids [43] could help improve corner detection quality, they are limited in actual engineering applications due to operational complexity or acceptance issues. One effective solution to solve this problem is introducing a hypothesize-and-test framework, i.e. RANSAC [44], to filter the negative influence of volatile features or corners for improving the reliability of calibration, which has been proven by existing literature from relative pose estimation [45] to intrinsics calibration [46]. Inspired by this, we propose calibrating the stereo imaging system by integrating Zhang's method into the RANSAC loop to provide a low operational cost method for achieving accurate and precise calibration in hand, thereby improving the reliability of subsequent ODS measurements in Section 2.2. Though this solution is not new, we here briefly introduce this practical calibration method for the sake of completeness of the proposed ODS sensing method.

We begin with a brief revisitation of the fundamental perspective projection model. Let $\mathbf{K} \in \mathbb{R}^{3\times3}$ be the internal parameter matrix of the perspective camera in our system, and $\mathbf{R}_j \in \mathrm{SO}(3), \mathbf{t}_j \in \mathbb{R}^3$ be the rotation matrix and translation vector of the camera relative to a given pose (j-th) of the planar target, on which the 3D corner coordinates are denoted by $\mathbf{X} \in \mathbb{R}^3$. Note that $\mathbb{R}^{3\times3}$ and $\mathbb{R}^3$ denote the sets of all 3×3 matrices and 3-dimensional vectors, respectively, and $\mathrm{SO}(3)$ is the special orthogonal group consisting of all 3D rotation matrices. Consider the assumption of perspective projection. The observed image point $\mathbf{u}$ of $\mathbf{X}$ in the pixel domain is given by the following model:

$$\mathbf{u} = \mathbf{K}D([\mathbf{R}_j|\mathbf{t}_j]\mathbf{X}), \tag{1}$$

where $D(\mathbf{x}) = (1 + k_1||\langle\mathbf{x}\rangle||^2 + k_2||\langle\mathbf{x}\rangle||^4)\langle\mathbf{x}\rangle$ is the nonlinear radial distortion function, defined by the first- and second-order radial coefficients k_1 and k_2, respectively, which maps an ideal image projection $\mathbf{x} = [\mathbf{R}_j|\mathbf{t}_j]\mathbf{X}$ to its distorted counterpart in the normalized image domain, $||\cdot||$ denotes the L_2-norm, and $\langle\cdot\rangle$ is an operator maps a 3-vector $[x, y, z]^T$ to $[x/z, y/z, 1]^T$. Further details on the distortion function can be found in [38,47]. According to Zhang's method [38], the camera observes the planar target from N different poses, resulting in N frames of calibration images correspondingly. Each corner on the planar target has N corresponding projections in the observed images, providing $2N$ constraints to determine the perspective projection mapping described in Eq. (1).

We evaluate the accuracy of the determined calibration parameters by using the average re-projection error in the pixel plane, given by:

$$e_j = \frac{1}{M}\sum_{i=1}^{M} ||\mathbf{u}_i - \mathbf{K}D([\mathbf{R}_j|\mathbf{t}_j]\mathbf{X}_i)||^2, \tag{2}$$

where the subscript i indexes the corner points with a total number of M. We estimate the optimal combination of $\mathbf{K}$ and $D(\cdot)$ by minimizing Eq. (2) within the framework of RANSAC. To describe the algorithm intuitively, we introduce two sets of points: one set $\{\mathbf{X}_i\}_{i=1}^M$ consists of global control points composed of 3D corners $\mathbf{X}_i$, and the other set $\{\{\mathbf{u}_j\}_{j=1}^N\}_{i=1}^M$ comprises their projections $\mathbf{u}_j$ observed from the N poses. Each element $\mathbf{X}_i$ in the control point set corresponds to N different elements $\{\mathbf{u}_j\}_{j=1}^N$ in the projection set, providing known 3D-2D correspondences between the target plane and each of the calibration images. By randomly selecting no fewer than four samples in the control point set, we perform Zhang's calibration to obtain a test result based on the correspondences between the two point sets. Subsequently, the re-projection errors Eq. (2) for all poses are averaged to determine the overall calibration error. If the calibration error meets the pre-defined threshold, the calibration parameters will be updated; otherwise, a subset of control points will be randomly selected again to perform Zhang's calibration until the error threshold or the sampling frequency is reached.

We experimentally found that the RANSAC-based calibration strategy could obviously improve the calibration accuracy in a straightforward yet simple manner (see Section 3.3), ensuring the precision of optical ODS sensing. It is worth noting that, in a stereo imaging system, the planar target is imaged by two or more cameras simultaneously for each calibration pose. The proposed RANSAC calibration method can be applied independently for each camera or jointly for both cameras. In the latter case, the corresponding projections of the control point samples in different cameras should correspond one-to-one. Both cases can equally determine the relative rotation matrix $\mathbf{R}$ and translation vector $\mathbf{t}$ between the two cameras. Furthermore, although each iteration in our RANSAC calibration procedure follows Zhang's method and theoretically requires only two poses of the planar target, practical considerations such as pose degeneracy necessitate using no fewer than eight image pairs (typically 8–25) to robustly estimate all calibration parameters, including the distortion coefficients.

2.2. Full-field ODS estimation in frequency-domain

Given the calibration parameters of the stereo imaging system and a set of stereo observations from successive frames of a structural object, it is possible to recover both the 3D shape and motion of the object in the time and frequency domains. To achieve this, we first introduce a single-parameter fundamental model through an improved stereo rectification strategy to link the object points to their stereo projections. Then, by incorporating the time dimension and image motion constraints, we develop a time-to-frequency transformation to characterize the ODS in the frequency domain.

Single-parameter fundamental model. Consider that the two cameras in our stereo imaging system have the same focal length. As shown in Fig. 2(a), the left and right cameras are referred to as $\mathbf{C}_1$ and $\mathbf{C}_2$, respectively, and the world frame is aligned with that of the left camera. With this configuration, we represent a 3D object point $\mathbf{X}$ by its depth parameter d through the back-projection as $\mathbf{X}(d) = \mathbf{x}_1 d$, with its normalized projection $\mathbf{x}_1 = \mathbf{K}^{-1}[\mathbf{u}_1^T, 1]^T$ and the image distortion has been removed according to D in Section 2.1. The depth parameterization lowers the parameters of reconstructing the object point $\mathbf{X}$ from three to one, formatting the ODS imaging as a single-parameter projective system. This implies that we could estimate the ODS by only computing the depth of each object point of interest (POI).

Therefore, we perform stereo rectification on the original imaging geometry in Fig. 2(a) by following the well-known Bouguet's algo-

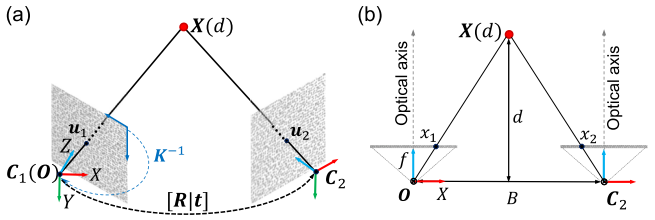

Fig. 2. Two-view imaging geometries with depth parameterization. (a) The original geometric configuration of the stereo vision system consists of left ($\mathbf{C}_1$) and right ($\mathbf{C}_2$) cameras, with the reference frame ($O-XYZ$) aligned to the left camera's coordinate system. This alignment allows mapping a pair of pixels ($\{\mathbf{u}_1, \mathbf{u}_2\}$) to a depth-parameterized 3D point $\mathbf{X}(d)$. (b) The rectified stereo geometry, in which the left and right image planes are mapped to a common plane at the ideal focal distance f, with parallel optical axes.

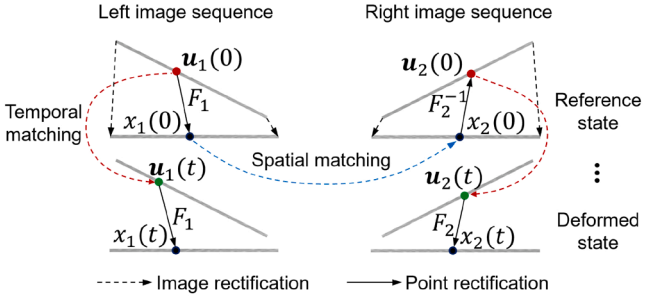

Fig. 3. Schematic diagram of spatial-temporal matching with local stereo rectification strategy. The stereo pair in the reference state is rectified using matrices F_1 and F_2. Each point of interest (POI), $\mathbf{u}_1(0)$, is mapped to rectified space as $x_1(0)$, and then matched to $x_2(0)$ using 1D epipolar-constrained spatial correlation matching. Its correspondence $\mathbf{u}_2(0)$ is recovered via inverse mapping using F_2^{-1}. Temporal matching is applied to track $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$, and their rectified coordinates $x_1(t)$ and $x_2(t)$ are determined at any deformed state. Only coordinates of POIs are rectified in the spatial-temporal matching process, avoiding interpolation of the full image.

rithm [48], obtaining a standard triangulation geometry with collinear epipolar lines, as shown in Fig. 2(b). Existing work has shown that errors introduced by stereo rectification could be acceptable for stereo matching and 3D deformation measurement tasks, even if the relative rotation between cameras is not in the same plane [49,50]. Additionally, stereo rectification involves geometric transformations of the image domain, followed by interpolation to obtain the grayscale values of displaced pixels, which introduces interpolation errors. To address this issue, we adopt a compromise strategy for stereo rectification, as shown in Fig. 3.

Specifically, we first rectify the stereo image pair obtained in the reference state using the stereo rectification matrices of the two cameras, denoted as F_1 and F_2, respectively. Meanwhile, mapping each POI $\mathbf{u}_1(0)$ defined on the original left reference image to its rectified position $x_1(0)$ and then, using spatial DIC matching under 1D epipolar constraints to find the corresponding point $x_2(0)$ in the rectified right image. By applying an inverse mapping with F_2^{-1}, we can obtain the stereo correspondence $\mathbf{u}_2(0)$ for the POI $\mathbf{u}_1(0)$. Subsequently, we perform temporal DIC matching to track $\mathbf{u}_1(0)$ and $\mathbf{u}_2(0)$ individually in the original left- and right-image sequences to find their positions $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$, respectively, in any deformed state. As a result, the rectified coordinates $x_1(t)$ and $x_2(t)$ of each stereo correspondence can be determined. In this way, stereo rectification is applied only to the coordinates of the displaced POIs rather than the entire deformed image pairs, thereby avoiding unnecessary image interpolation. After rectification, we can estimate the depth parameter according to the standard triangulation geometry in Fig. 2(b), as follows:

$$d = \frac{Bf}{\delta}, \tag{3}$$

where B and f are the baseline width and the ideal focal length, respectively, and δ, the disparity of viewing $\mathbf{X}(d)$ from the rectified stereo view, is computed as $\delta = x_2 - x_1$. We can then derive the fundamental model that describes the object point $\mathbf{X}$ in relation to the disparity in the following form:

$$\mathbf{X}(\delta) = Bf\frac{\mathbf{x}_1}{\delta}. \tag{4}$$

This formula is built upon the depth parameterization, but in mathematical form, we follow it up by expressing it as a function of disparity since it is more straightforward to represent the ODS in the frequency domain. This will be shown in the following content.

Frequency-domain perspective transformation. Following the assumption in continuum mechanics that chooses the undeformed configuration as the reference, we can describe the vibration of any observed object point by tracking its positional changes in a Newtonian reference frame from the measured temporal stereo correspondences. This means, at any sampling time t, the deformed coordinates $\mathbf{X}(t)$ of the object point can be expressed relative to the reference position $\mathbf{X}_0$ by using the Lagrangian description of motion as:

$$\mathbf{X}(t) = \mathbf{U}(t) + \mathbf{X}_0, \tag{5}$$

where $\mathbf{U}(t)$ is the displacement vector indicating the motion information in 3D Euclidean space.

With the fundamental coordinate formation of the object point in Eq. (4), we have a choice of writing the displacement vector at the time t with respect to the disparity δ in the temporal stereo image domain

$$\mathbf{U}(t) = Bf\left(\frac{\mathbf{x}_1(t)}{\delta(t)} - \frac{\mathbf{x}_1(0)}{\delta(0)}\right), \tag{6}$$

where $\mathbf{x}(t)$ and $\mathbf{x}(0)$ denote the current deformed image coordinates and their counterparts in the reference image. Eq. (6) allows for a fast and efficient solution strategy for the 3D displacement with a few algebraic operations if we approach the spatial-temporal stereo tracking in the image plane.

To derive the ODS of a specific frequency from Eq. (6), we consider a multiple-DOF linear time-invariant structural system. The forced vibration equations are given by:

$$\mathbf{M}\frac{\partial^2\mathbf{U}(t)}{\partial t^2} + \mathbf{C}\frac{\partial\mathbf{U}(t)}{\partial t} + \mathbf{S}\mathbf{U}(t) = \mathbf{\Phi}(t) \tag{7}$$

where $\mathbf{M}, \mathbf{C}, \mathbf{S}$ are the constant coefficient matrices of mass, damping, and stiffness, respectively, and $\mathbf{\Phi}(t)$ is the exciting force vector for all DOFs. Consequently, by performing discrete Fourier transformation to Eq. (7), we can obtain an equivalent formulation in the frequency domain as follows:

$$\mathbf{U}(\omega) = \mathbf{H}(\omega)\mathbf{\Phi}(\omega), \tag{8}$$

where ω denotes the frequency parameter and $\mathbf{H}(\omega)$ is a complex matrix representation of the frequency response function [51]. Every element of $\mathbf{H}(\omega)$ represents the frequency response relationship of one DOF to another, including amplitude and phase information. It is worth noting that the frequency response matrix can be decomposed into a sum of multiple matrices, each representing the contribution of a single vibration mode for harmonic excitation [52]. This implies that the response $\mathbf{U}(\omega)$ constitutes the vibration modes of the system at the frequency ω, providing a fundamental support for measuring the full-field ODS $\mathbf{U}(\omega)$ through synchronous stereo point motion as shown in Eq. (6). Furthermore, the synchronous 3D imaging of structural vibration ensures that the motion of object points is measured from different views without delay, thereby eliminating phase errors in the observed response compared with the existing method [6].

Based on the fundamental observations above, we can derive an efficient visual ODS representation in the complex domain by applying the Fourier transformation to Eq. (6) over the imaging time range $[0, T]$ as

follows:

$$\mathbf{U}(\omega) = Bf \int_0^T \left(\frac{\mathbf{x}_1(t)}{\delta(t)} - \frac{\mathbf{x}_1(0)}{\delta(0)} \right) e^{-i2\pi\omega t} dt. \tag{9}$$

For notation simplicity, we introduce a state vector:

$$\boldsymbol{\xi}(t) = \frac{\mathbf{x}_1(t)}{\delta(t)} \in \mathbb{R}^3, \tag{10}$$

which gives a compact stereo image motion encoding by combining the dynamic image of an object point and its disparity jointly. Consequently, we have a simple linear mapping between the image motion state and the ODS response in the frequency domain:

$$\mathbf{U}(\omega) = Bf \int_0^T (\boldsymbol{\xi}(t) - \boldsymbol{\xi}(0)) e^{-i2\pi\omega t} = Bf \cdot \Delta\boldsymbol{\xi}(\omega), \tag{11}$$

where $\Delta\boldsymbol{\xi}(\omega)$ can be regarded as the disparity-weighted relative image displacement amplitude in the frequency domain.

Based on the depth parameter model represented by disparity in Eq. (6), we can observe that the state vector $\boldsymbol{\xi}(t)$ effectively encapsulates the motion of the stereo images into a compact dynamic vector, together with a scalar variable, namely the time-varying $\delta(t)$, to represent the time-varying normalized 3D coordinates of the object points, with a maximum magnitude of 1. Therefore, Eq. (9) can be reformulated as the Fourier transform of the relative changes of the state vector, allowing the displacement amplitude in the frequency domain (i.e., the ODS response) to be expressed as a linear function of the state vector changes $\Delta\boldsymbol{\xi}(\omega)$ in the frequency domain without introducing tradition triangulation process. In this way, Eq. (11) forwards the problem of ODS estimation to one of performing the Fourier transformation to temporal changes of the weighted state vector in the normalized image domain, significantly reducing the complexity compared to the existing methods. What followed is that we can compute the current frequency-domain spatial coordinates $\mathbf{X}(\omega)$ by transforming $\Delta\boldsymbol{\xi}(\omega)$ to its reference position (corresponding to $t = 0$), i.e., $\mathbf{X}(\omega) = \mathbf{U}(\omega) + Bf\boldsymbol{\xi}(\omega)_{t=0}$, leading to

$$\mathbf{X}(\omega) = Bf\boldsymbol{\xi}(\omega), \tag{12}$$

with $\boldsymbol{\xi}(\omega) = \Delta\boldsymbol{\xi}(\omega) + \boldsymbol{\xi}(\omega)_{t=0}$, and the reference state vector $\boldsymbol{\xi}(\omega)_{t=0}$ is determined according to the selection of the reference configuration.

Estimation of the ODS response. The algorithm of ODS estimation for a calibrated stereo imaging system can be summarized as follows:

- Perform stereo rectification in the reference state in Fig. 2, so that the relative rotation matrix $\mathbf{R} = \mathbf{I}$ and translation vector $\mathbf{t} = [B, 0, 0]^T$ to simplify the stereo matching complexity for building the spatial-temporal correspondences.
- Build stereo correspondences by applying spatial DIC matching across the rectified stereo pair in the reference state. After that, we perform temporal DIC tracking to track the deformed stereo positions $\{\mathbf{u}_1 \leftrightarrow \mathbf{u}_2\}$ and their rectified coordinates according to Fig. 3.
- Evaluate the state vector series $\{\boldsymbol{\xi}(t)\}_{t=0}^T$ using the tracked stereo correspondences by following Eq. (10).
- Estimate the ODS field $\mathbf{U}(\omega)$. As the state vectors are sampled discretely, their frequency-domain response $\Delta\boldsymbol{\xi}(\omega)$ can be readily obtained by applying the discrete version of the Fourier transformation in Eq. (11). Subsequently, the ODS $\mathbf{U}(\omega)$ at the frequency ω can be estimated directly according to Eq. (11). Meanwhile, the estimated ODS fields at all frequency components constitute the response spectrum in the frequency domain.

2.3. Damage identification based on ODS measures

After obtaining the ODS $\mathbf{U}(\omega)$, it becomes possible to extract the vibration modes for various feature inspection tasks, such as damage identification. To accomplish this, we introduce wavelet transformation [53] which is capable of performing multi-scale analysis on the original ODS signal. Although the ODS in the frequency domain is associated with 3D points on the object surface, it can be tracked one-to-one in the left reference image frame without projecting the reconstructed ODS back to the image plane, since each point has a well-defined projection in the left image according to the imaging geometry shown in Fig. 2. This allows us to build a manifold representation of the ODS in the left reference image domain (at the reference state of $t = 0$) with 2D grid coordinates $\mathbf{u}$. For the purposes of subsequent formula representation, we denote the ODS manifold as $\mathbf{U}(\mathbf{u}; \omega)$ to show the ODS response at a given frequency ω varies with the reference image coordinates $\mathbf{u}$. (For the sake of brevity, we here omit the time subscript of $\mathbf{u}$.) Because of the bijective relationship described in Eqs. (4) and (6), the manifold $\mathbf{U}(\mathbf{u}; \omega)$ is highly structured to allow us to query and transform the ODS data by tracking the image coordinates only, enabling damage identification by applying a 2D wavelet transformation on the 2D grid as follows, rather than a 3D transformation on the 3D point cloud.

Giving that a mother wavelet $\psi(\mathbf{u}) \in \mathcal{L}(\mathbb{R}^2)$ relative to the image domain, we can obtain a family of wavelet functions by transforming $\psi(\mathbf{u})$ using shift parameters $\mathbf{a} \in \mathbb{R}^2$ and a scale factor $s > 0$:

$$\psi(\mathbf{u}; \mathbf{a}, s) = \frac{1}{s}\psi\left(\frac{\mathbf{u} - \mathbf{a}}{s}\right), \tag{13}$$

where $\mathcal{L}$ denotes a square-integrable Hilbert space defined on the image domain. With this mother wavelet definition, the continuous wavelet transformation of signal $\mathbf{U}(\mathbf{u}; \omega)$ is given by

$$\mathcal{W}(\omega; \mathbf{a}, s) = \frac{1}{s} \int \mathbf{U}(\mathbf{u}; \omega)\psi^*\left(\frac{\mathbf{u} - \mathbf{a}}{s}\right) d\mathbf{u}, \tag{14}$$

where $\psi^*(\cdot)$ stands for the complex conjugate counterpart of the wavelet in Eq. (13). The resulted wavelet coefficients $\mathcal{W}(\omega; \mathbf{a}, s)$ encode the damage information due to stiffness loss. In practice, as $\mathbf{U}(\omega)$ is measured associating to the discrete image domain $\mathbf{u}$, Eq. (14) is often approached as a matrix transformation by choosing a suitable wavelet function to identify the damage parameters in a blend of mathematical elegance and practical ingenuity according to the 2D wavelet transformation on ODS manifold.

At this point, we regard the ODS fields as images to enable the matrix-based 2D wavelet transformation by introducing the well-known Haar wavelet, the simplest orthogonal wavelet, whose rectangular basis functions are well-suited for our purpose of identifying both the location and degree of damage without imposing specific limitations on the shapes or orientations of the damage. The Haar transform is based on 1D rectangular-shaped Haar functions, denoted by $\psi_n(x)$, that are defined over the continuous interval $x \in [0, 1)$, where $n \in \mathbb{R}^+$ can be decomposed as $n = 2^p + q$ uniquely. The parameters p and q determine the amplitude ($2^{p/2}$) and width ($\frac{1}{2^{p+1}}$) of the rectangular waves of $\psi_n(x)$ and their locations ($\frac{q+0.5}{2^p}$) along x, respectively. The amplitude parameter significantly affects the strength of the Haar wavelet transform and the width parameter determines the coverage of the wavelet in the time domain and consequently affects the frequency resolution of the ODS signal. It can be found that a larger value of p results in a higher amplitude but a narrower width. This also means that noise is amplified, and frequency resolution is sacrificed. Taking this into consideration, we experimentally adopt a compromise based on the characteristics of the ODS signal: using the Haar mother wavelet function without translation or scaling, i.e., setting $p = q = 0$ and thus $n = 1$, generating the Haar mother wavelet function as follows:

$$\psi(x) = \begin{cases} 1, & 0 \le x < 0.5 \\ -1, & 0.5 \le x < 1 \\ 0, & \text{elsewhere} \end{cases} \tag{15}$$

Compared with smoother wavelets such as Morlet and Daubechies, the Haar wavelet offers sharper spatial localization and avoids spreading local anomalies over larger regions by specializing $p = q = 0$. Morlet wavelets, with their oscillatory and Gaussian-modulated structure, prioritize frequency resolution but are less effective at highlighting sudden changes. Daubechies wavelets, though compactly supported, are

smoother and distribute abrupt variations over multiple coefficients, which could weaken the visibility of localized damage in ODS fields. This choice provides a balanced compromise: it maintains strong sensitivity to local discontinuities essential for damage localization, minimizes noise amplification by avoiding high-amplitude scales with $p = 0$ (unit scale), and ensures that the frequency resolution remains sufficient for distinguishing ODS anomalies. The resulting mother wavelet in Eq. (15) delivers both computational efficiency and reliable detection performance under practical noise conditions.

Giving a father scaling function $\varphi(x)$ with the constant value 1 on $[0, 1)$ that obeys the fundamental requirements of multiresolution analysis [54], there exists a discrete form of the mother wavelet function in Eq. (15) with integer shifts k and binary scalings 2^j

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k), \tag{16}$$

where j is the decomposition level which maximum value is determined by the signal length N in the leading dimension of the image domain equipped to the ODS manifold. The formula shows that each $\psi_{j,k}$ is obtained from ψ through the discrete samplings of the shift parameter and the scale factor. Noting that, after the ODS signal is decomposed via Eq. (16), its length at the j-th level is reduced to $N/2^j$ by downsampling. This means that the downsampling operation is equivalent to adjusting the wavelet translation step size to 2^j, thereby automatically generating a sequence of shift parameters as $k = 0, 1, \ldots, N/2^j - 1$. It can be seen that as long as specifying a suitable decomposition level $j \in [0, j_{\max}]$, the discrete Haar mother wavelet can be completely defined.

Considering the scale parameter determines the frequency band range of the wavelet analysis, the relation between the decomposition level and the frequency should be satisfies Nyquist's sampling law: $\omega_c = \omega_s/2^{j+1}$, with ω_s the signal sampling frequency and ω_c the center frequency of the wavelet band. Since the damage vibration response appears as a step signal on the ODS manifold, the convolution between the step signal and the wavelet window reaches an extreme value–indicating the damage signal–when they are aligned well. Therefore, the appropriate range for j can be inversely determined by setting $\omega_c \approx \omega_d$ [55], yielding $j = \log_2(\omega_s/2\omega_d)$, where ω_d, the spectral frequency component corresponding to the damage response due to stiffness loss, is determined from the actual ODS spectrum.

As a result, we can build $N \times N$ Haar wavelet transform matrix $\mathbf{W}$ based on Eqs. (15) and (16), as shown in the following concise recursive equations:

$$\mathbf{W}_N = \begin{bmatrix} \mathbf{W}_{N-1} \otimes [1, 1]^T & \mathbf{I}_{2^{N-1}} \otimes [1, -1]^T \end{bmatrix} \tag{17}$$

with $\mathbf{W}_0 = [1]$, where $\otimes$ denotes the Kronecker product (or tensor product) and $\mathbf{I}$ is an identity matrix which has a compatible shape with $\mathbf{W}$. To compute the wavelet coefficients $\mathcal{W}$ for the ODS manifolds $\mathbf{U}(\mathbf{u}; \omega)$, we first map the rows of $\mathbf{U}$ to their Haar coefficients using the matrix $\mathbf{W}_N^{-1}$ and then convert the resulted temporal coefficient matrix along the column coordinate using $\mathbf{W}_M^{-1}$. Because the rows and columns are exchanged in the first step, we can finally obtain the wavelet coefficients $\mathcal{W}$ as

$$\mathcal{W} = \langle \mathbf{W}_M^{-1} \mathbf{U} \left(\mathbf{W}_N^{-1}\right)^T \rangle, \tag{18}$$

where $\langle \cdot \rangle$ is the normalization operator. It should be noted that here M and N are the dimensions of the gridded region of interest (denoted as ROI for short) in which the interest points for computing the displacement and ODS are defined, i.e., $\mathbf{u} \in \mathbb{R}^{M \times N}$ for $\mathbf{U}(\mathbf{u}; \omega)$. Once we obtained the normalized wavelet coefficients, the damage location and significance can be identified via maximum detection with the aid of the non-maximal suppression strategy optionally. As the ROI is usually a user-defined region within the left reference image, a simple transformation is required to map the ROI coordinates of the identified damage to world coordinates. Suppose the specified origin of the ROI is $\mathbf{u}_o$; the coordinates of each ROI point in the reference image are given by $\mathbf{u}_1' = \mathbf{u} + \mathbf{u}_o$. We can then find its correspondence $\mathbf{u}_2'$ in the right image by following the rectified spatial matching shown in Fig. 3, allowing computation of the corresponding 3D world coordinates using Eq. (4) or via trivial triangulation with Eq. (1), regardless of whether the underlying surface is planar or curved. In this process, the uncertainty is consistent with that of 3D reconstruction, mainly arising from calibration and spatial matching, as discussed in Section 3.3. Once the world coordinates of each identified damage are obtained, its location in the object frame can be determined (a possible solution is presented in Section 2.3).

It is worth noting that, in this work, we identify the damage significance from ODS by following the linear time-invariant vibration system in Eq. (7). Since the ODS at each frequency characterizes the amplitude response of a single degree of freedom in the vibrating system, basic principles of structural dynamics imply an approximately inverse relationship between ODS and stiffness as $E \approx \frac{|\Phi_0|}{|\mathbf{U}|}$, where $|\Phi_0|$ and $|\mathbf{U}|$ indicates the amplitudes of the initial excitation and the detected ODS response, respectively. According to the linear isotropic damage model proposed in [56], the damage significance, denoted by ϕ, can be defined in terms of the stiffness as:

$$\phi = 1 - \frac{E_d}{E_0}, \tag{19}$$

where E_d and E_0 denote the damaged and initial stiffness, respectively, which can be theoretically estimated using the equivalent stiffness definition for beams in elementary mechanics of materials, given the known cross-sectional geometry of the structural component. Therefore, we could represent the damage significance as a local relative change in the amplitude of ODS, reflecting the corresponding local reduction in stiffness, as follows:

$$\phi \approx 1 - \frac{|\mathbf{U}|}{|\mathbf{U}_d|}, \tag{20}$$

where $|\mathbf{U}_d|$ denotes the ODS amplitude at damaged location(s), allowing us to identify the damages via the normalized wavelet coefficients in Eq. (18).

3. Experiments and results

3.1. Experimental setup

In accordance with the stereo imaging geometry in Fig. 2, we configured an experimental setup with two high-speed cameras (Revealer G536M Pro,1.1″ monochrome CMOS sensor, 3100 fps, minimum exposure time 500 μs, global shutter) as shown in Fig. 4, where the left and right cameras are labeled. These two cameras are identical with a resolution of 2560×1920 pixels and were equipped with 24 mm lenses (AF-S NIKKOR, f/1.8G). Though the two cameras were mounted independently due to the limitation of measurement conditions in Fig. 4, it is preferred that the cameras be mounted on the same crossbar to provide

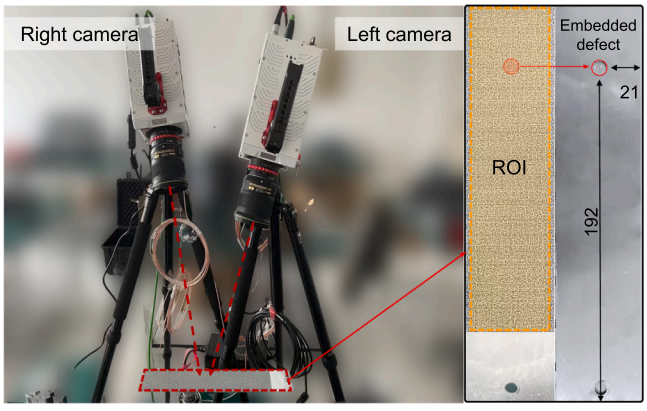

Fig. 4. Experimental setup for ODS and damage identification. The setup consists of two high-speed cameras (labeled by left and right camera, respectively) and a damaged specimen, with the location of the embedded defect measured in millimeters.

a relative rotation angle in one plane to simplify the stereo rectification. The specimen used is a 1.5 *mm* thick metal component made of 304 stainless steel, with dimensions of 300 *mm* in length and 50 *mm* in width. The flatness of the specimen is ±5 μm. It is worth noting that the specimen is recommended to first fabricate a blind hole on a thicker base material and then use a precision surface grinding machine for shaping to ensure the flatness. To validate the performance of our damage detection method, we prefabricated a blind hole as a defect to be identified. This defect has a depth of 1.0 *mm* and a diameter of 8.0 *mm*, located on one side of the specimen as shown in the right part of Fig. 4. The other surface was coated with a random speckle pattern using hydrographic transfer printing to facilitate the DIC matching technique, as the raw surface pattern of the specimen was textureless. The specimen was mounted on a vibration testing platform through a mounting hole and was excited by a JZ-1 electromagnetic exciter installed on the platform. The excitation point was located in the middle of the specimen. Additionally, a 1 KHz magnetic displacement transducer (DT), with a sensitivity of 280 $mv/cm/s$ and an accuracy of ±1 μm, was attached at the location of the defect to read the displacement data via a dynamic signal acquisition and processing system.

3.2. Measurement details

In conducting the test, the specimen was first brought to a state of equilibrium by adjusting the stinger of the exciter. Subsequently, the stereo imaging system was calibrated using the proposed calibration method in Section 2.1 with a planar target comprising 12×9 checkerboard pattern. The calibration target was fabricated from ceramic material with geometric dimensions of 200 $mm \times 160$ $mm \times 4$ *mm* and a pattern size of 14 $mm \times 14$ *mm*. Additionally, a test planar target featuring an 8×8 checkerboard grid was employed to validate the accuracy of the imaging system calibration from a 3D reconstruction perspective. This test target was made of transparent float glass with geometric dimensions of 150 $mm \times 150$ $mm \times 2$ *mm* and a pattern size of 11 $mm \times 11$ *mm*. Given that the manufacturing precision and flatness of the calibration and test targets are critical to measurement accuracy, both were custom-made with a pattern size tolerance of ±1 μm and a flatness of ±2 μm. After calibration, the stereo camera system was triggered synchronously to capture a pair of speckle images as a reference stereo for subsequent computation. It is worth noting that, in the calibration process, 18 pairs of calibration images were synchronously captured for camera calibration. In addition, we supposed that 10 % to 30 % of corners in each calibration image were outliers for sampling the appropriate corners to determine the parameters. The calibrated internal parameters of both cameras are listed in Table 1 and the relative external parameters between them are listed in Table 2. For validation purposes, the calibration parameters obtained by Zhang's method are also presented. Note that, by evaluating the distortion function D in Eq. (1) using the calibrated distortion coefficients for both cameras, we found that the distortion functions estimated by our method are consistent with those obtained by Zhang's method in terms of typical lens distortion, but exhibit a slightly lower overall distortion level. As shown in Table 2, the re-projection error of our method is lower than that of the original Zhang's method (the errors of both methods were computed with all calibration image pairs without removing the ones with large errors), indicating that the proposed algorithm in Section 2.1 achieves an expected calibration performance.

Table 1
Calibrated internal parameters.

Camera	Parameter	Zhang's [57]	This work
C_1	f_x, f_y	2799.43, 2800.09	2802.78, 2803.09
	c_x, c_y	1281.31, 1017.01	1282.21, 1012.69
	k_1, k_2	-0.068, 0.182	-0.079, 0.180
C_2	f_x, f_y	2791.06, 2795.13	2790.24, 2792.85
	c_x, c_y	1258.80, 1029.56	1290.73, 1001.68
	k_1, k_2	-0.064, 0.207	-0.012, 0.671

Table 2
Calibrated relative external parameters.

Method	Rotation angle (°)	Translation vector (mm)	Error
This work	(-10.89, 23.33, 0.42)	(-158.95, -62.32, 46.27)	0.039
Zhang's [57]	(-9.71, 23.35, 1.14)	(-155.52, -56.07, 35.63)	0.088

After recording the reference configuration, the exciter applied continuous excitation to the specimen at 17 Hz, close to its first natural frequency. Note that, since this specimen is used for comprehensive performance verification of the proposed method in subsequent Section 3.4, this excitation ensures that the system operates near resonance, producing sufficiently large amplitudes while maintaining the camera's depth of field. This setup provides an adequate field of view and avoids additional DIC matching errors caused by defocus and small displacement to balance accuracy validation and optical constraints. During this period, the vibration of the specimen was observed synchronously by the stereo cameras at a frame rate of 1000 fps. Limited by the built-in memory of the cameras, the entire test procedure lasted five seconds, resulting in a total of 5000 stereo pairs. To estimate the ODS, we defined a ROI in the left reference image as the computing domain, which is indicated by the yellow rectangle in Fig. 4.

3.3. Evaluation of stereo calibration method

In addition to the re-projection error that is a necessary condition, we further evaluated the camera calibration results by computing the reconstruction errors for both the planar calibration and test targets using the calibration parameters in Tables 1 and 2. To this end, we define a world coordinate frame on each of the planar targets where the origin aligns with the bottom-right corner point and the X and Y axes point to the right and top, respectively; the Z axis is thus perpendicular to the target surface according to the right-hand rule. With this coordinate system definition, we can determine the 3D coordinates of all corners on each target as the ground truth (GT) for comparing the reconstruction error of our method with that of the original Zhang's method. Then, both targets were imaged to obtain two sets of 108 and 64 pixel correspondences of 3D corner points, respectively. By applying the calibration parameters obtained by our method and Zhang's method for triangulation, we can obtain two sets of reconstructed 3D points for each planar target. The 3D reconstruction results for the 12×9 calibration target and the 8×8 test target are shown in Fig. 5(a) and (b), respectively. In comparison, for both planar targets, the 3D corner points obtained by our RANSAC-based calibration method are closer to the GT plane (e.g., $Z = 0$).

For quantitative verification, we computed the reconstruction errors by following the 3D Euclidean distance metric $||\mathbf{X} - \mathbf{X}_{GT}||$ for the corner reconstructions in Fig. 5(a) and (b). The reconstruction error distributions for Fig. 5(a) and (b) are presented in Fig. 5(c) and (d), respectively, and the maximum and minimum values, mean and standard deviation (SD) of the reconstruction error were evaluated in each case as listed in Table 3. For the 12×9 calibration target, Zhang's method exhibits an error range of 0.010 to 0.230 *mm*, with a mean and an SD of 0.085 and 0.042 *mm*, respectively. In contrast, the RANSAC-based method significantly reduces these errors, achieving a narrow error range of $[0.002, 0.029]$ *mm* with a lower mean of 0.014 *mm* and an SD 0.006 *mm* – an order of magnitude lower than that of Zhang's method. Similarly, for the 8×8 test target, our RANSAC-based method again demonstrates superior performance, with a reduced error range of 0.003 to 0.025 *mm*, a lower mean of 0.013 *mm* and an SD of 0.005 *mm*, than the conventional Zhang's method, whose error distribution ranges from 0.052 to 0.292 *mm* with a significantly larger SD 0.043 *mm*. These results highlight the effectiveness of the RANSAC-based calibration method in minimizing reconstruction errors. Notably, the SD of the reconstruction errors for both

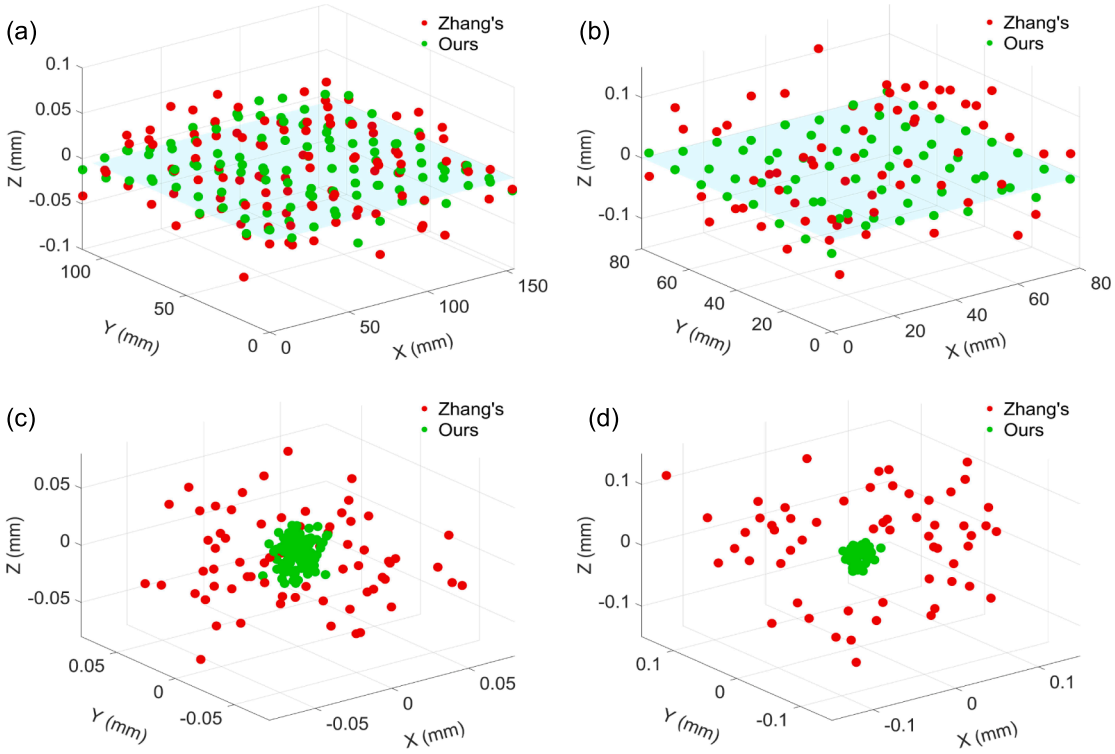

Fig. 5. 3D reconstruction results of the corner points. (a) The planar calibration target and (b) the test target using the conventional Zhang's method and the proposed RANSAC-based calibration method. The error distributions for the 3D reconstructions in (a) are shown in (c), while those for (b) are presented in (d).

Table 3
Comparison of 3D reconstruction errors between the regular Zhang's method and the RANSAC-based calibration method in this work.

Target	Method	3D reconstruction error (mm)			
		Maximum	Minimum	Mean	SD
12×9	Zhang's [57]	0.230	0.010	0.085	0.042
	This work	0.029	0.002	**0.014**	**0.006**
8×8	Zhang's [57]	0.292	0.052	0.124	0.043
	This work	0.025	0.003	**0.013**	**0.005**

targets indicates that the RANSAC-based calibration method achieves an order-of-magnitude improvement in the precision of 3D reconstruction, which is crucial for the stereo imaging system to capture small ODS responses and further identify high-order mode shapes. We attribute the superior performance to the ability of the RANSAC-based calibration to robustly reject outlier corners caused by noise, defocus, or uneven illumination, relying only on consistent points to estimate the calibration parameters with high stability. In contrast, the original Zhang's method relies solely on L_2-based optimization of the re-projection error in Eq. (2), which is highly sensitive to even a few inaccurate corners, reducing both accuracy and stability.

Subsequently, we investigate the performance of the proposed ODS estimation method by reconstructing the depth map for the reference configuration since it indicates the quality of the disparity map, which determines the correctness and accuracy of ODS estimation, as shown in Section 2.2. To this end, we estimated the disparity map by following the first three steps. The resulting disparity map is shown in Fig. 6(a). Subsequently, the depth for each observed point was reconstructed by applying Eq. (3), generating a depth map as shown in Fig. 6(b). In terms of appearance, both the disparity and depth maps are smooth, implying that the observed object points lie on a flat surface, which corresponds well to the actual geometry of the specimen.

In order to quantify the reconstruction error, we compared the estimated depth map of our method with that reconstructed according to Zhang's method, where the benchmark is an ideal plane. This is reasonable because the specimen surface used in the experiment is planar with plateness of $\pm 5\ \mu m$. The depth error maps resulting from the calibration results of the RANSAC-based method and Zhang's method are shown in Fig. 6(c) and (d), respectively. Fig. 6(c) shows that the depth reconstruction error induced by our RANSAC-based calibration method ranges between $-0.06\ mm$ and $0.09\ mm$, with a mean value of $0.02\ mm$; while in Fig. 6(d), the depth error induced by Zhang's calibration method ranges between $-0.17\ mm$ and $0.35\ mm$, with a mean value of $0.09\ mm$, which is significantly higher than that in (c). This comparison result shows that the RANSAC-based calibration method gains higher accuracy as we expected. Although this error level does not directly represent the reconstruction accuracy of spatial points, it indicates the proposed method's performance in recovering the relative 3D shape and thus, can be used to validate the accuracy and reliability of estimating ODS, which is also a relative geometry as given in Eq. (11).

Finally, we evaluated the computational efficiency of the proposed calibration algorithm,which involves three parameters: the inlier threshold, the number of iterations, and the calibration error. However, the number of iterations is determined by the inlier threshold for a given target calibration error. In vision-based deformation measurement tasks, a practical benchmark for calibration error is 0.04 pixels, which is widely adopted as the default metric for stereo vision system calibration. Therefore, in this evaluation, the only parameter that needs to be specified is the inlier rate, or equivalently, the outlier rate that should be filtered during RANSAC loop. Based on the configuration in Section 3.2, where 10 % to 30 % of the corners in each calibration image are considered outliers, computational efficiency was evaluated using inlier thresholds of 70 %, 80 %, and 90 %.

Table 4 summarizes the computational efficiency of the proposed RANSAC-based calibration under different inlier thresholds. The results show that increasing the inlier threshold substantially reduces the num-

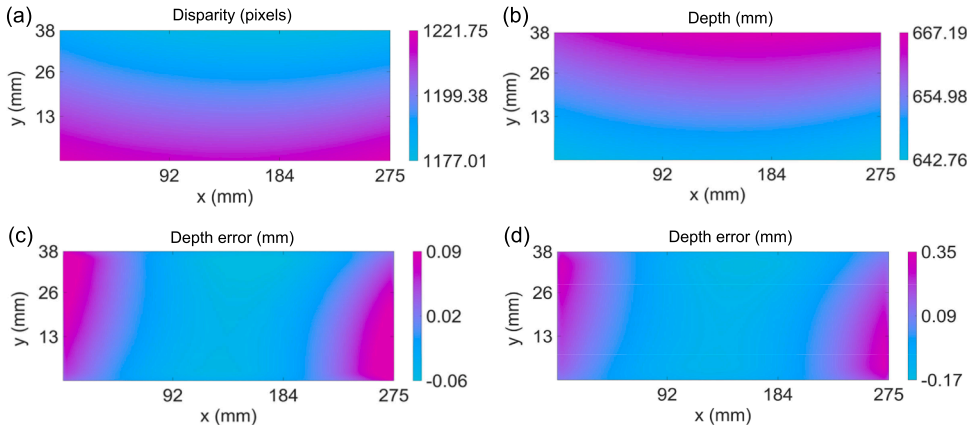

Fig. 6. Performance evaluation of the proposed calibration method in terms of disparity and depth estimation. (a) disparity map and (b) depth map estimated according to our calibration in this work, (c) and (d) are the depth error maps corresponding to our proposed calibration method and Zhang's method, respectively.

Table 4
Computational efficiency evaluation of our RANSAC-based calibration algorithm with respect to inlier threshold for a given target calibration error 0.04 pixels.

Inlier threshold	Number of iterations	Elapsed time (s)
70 %	330	5.92
80 %	64	3.84
90 %	13	2.80

ber of iterations and the elapsed time: 330 iterations (5.92 s) for 70 %, 64 iterations (3.84 s) for 80 %, and 13 iterations (2.8 s) for 90 %. It should be noted that the reported running times are dependent on the computer configuration. We chose an Intel Core I7-12700K CPU with 32 GB DDR4 memory in our implementation and the elapsed times are provided here for reference.

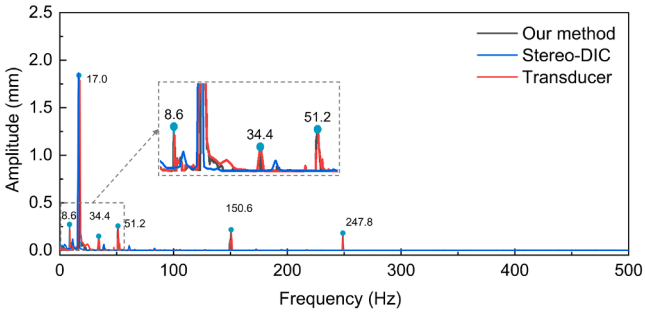

Fig. 7. Comparison of the measured ODS spectra. Comparison of the ODS spectra obtained using the proposed method and the conventional stereo digital image correlation (stereo-DIC) method, relative to the reference spectra measured by a displacement transducer.

3.4. Performance validation of ODS and damage sensing

Demonstration of ODS sensing. With the procedures outlined in Section 2.2, we estimated the ODS fields within the ROI in Fig. 4 for various frequencies. To show the correctness and feasibility of our method, we first investigated the resulting ODS spectrum corresponding to the position of DT and compared it with the corresponding spectrum obtained from the data recorded by the DT, as shown in Fig. 7. In addition, we calculated the ODS spectrum curve at this point using the 3D displacement measured by stereo-DIC, a representative of full-field deformation measurement methods based on optical triangulation, as shown by the blue curve in Fig. 7. The spectrum curves in the range of 0 to 51.2 Hz are zoomed in for a detailed comparison. Let the ODS spectrum of the DT be the benchmark. Both our method and the stereo-DIC method successfully capture the ODS response at the resonant frequency of 17 Hz. The corresponding amplitudes are 1.84 *mm* and 1.80 *mm*, respectively, compared with the GT value of 1.86 mm, yielding absolute errors of 0.02 mm and 0.06 mm (relative errors of 1.1 % and 3.2 %), respectively. For other frequency components, the stereo-DIC-based method failed to capture most of the ODS amplitudes, whereas our method was able to capture the complete spectrum. Even for the frequency at 34.4 Hz with the smallest amplitude of 0.14 *mm*, the estimation error remains only 0.01 *mm*. We attribute the main reason to the noisy spatio-temporal displacement reconstructed by stereo-DIC cannot be cleanly separated from noise after transformation into the frequency domain, requiring local filtering that inevitably suppresses high-frequency motion content. In contrast, our method separates vibration signals from the noise directly through frequency-domain triangulation, enabling high-fidelity recovery of ODS responses across the entire frequency range. These results in Fig. 7 demonstrate that our method reliably captures the correct dynamic response even for vibration amplitudes as small as 0.14 *mm*, showing substantially higher robustness than stereo-DIC to noise and amplitude variations.

As a demonstration of the measured ODS fields, we focus on specific peaks of the ODS amplitude curve marked in Fig. 7. Though the frequency coordinates at which the ODS response spikes are clearly observable, it is important to note that each point in the inset spectrum figure represents an ODS field. Therefore, we illustrate four ODS fields containing points at 17.00 Hz, 51.20 Hz, 150.60 Hz, and 248.80 Hz, respectively, as shown in Fig. 8. It is evident that the operating deflection amplitude decreases with increasing vibration frequency, while the shape and structure become more complex. This phenomenon precisely aligns with the theoretical trend of the displacement amplitude response of the component used in our experiment. It is worth noting that, although the accuracy of ODS sensing is validated at a single point using the displacement transducer, the accuracy at additional spatial points is expected to follow the same vision-based geometric consistency. This is because the ODS reconstruction is derived from the linear mapping in Eq. (11) under a pre-calibrated stereo vision geometry, which provides a fixed and consistent global constraint on the ODS estimation for all object points of interest.

Considering that we assume the ODS represents the structural response by linearly encapsulating the vibration contributions, we can identify the mode shapes associated with the detected frequencies in Fig. 7 from the sensed ODS fields with singular value decomposi-

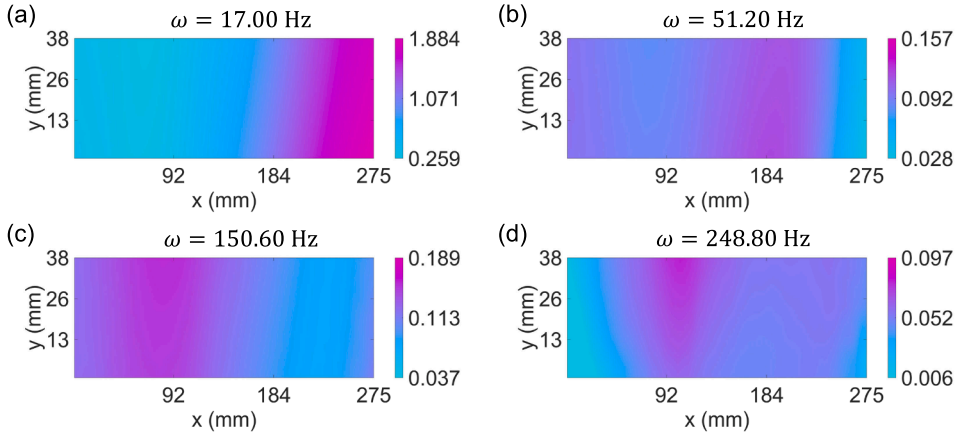

Fig. 8. Demonstration of the measured ODS fields. Visualization of the full-field ODS response of the structure at representative frequencies: (a) 17.0 Hz, (b) 51.2 Hz, (c) 150.6 Hz, and (d) 248.8 Hz.

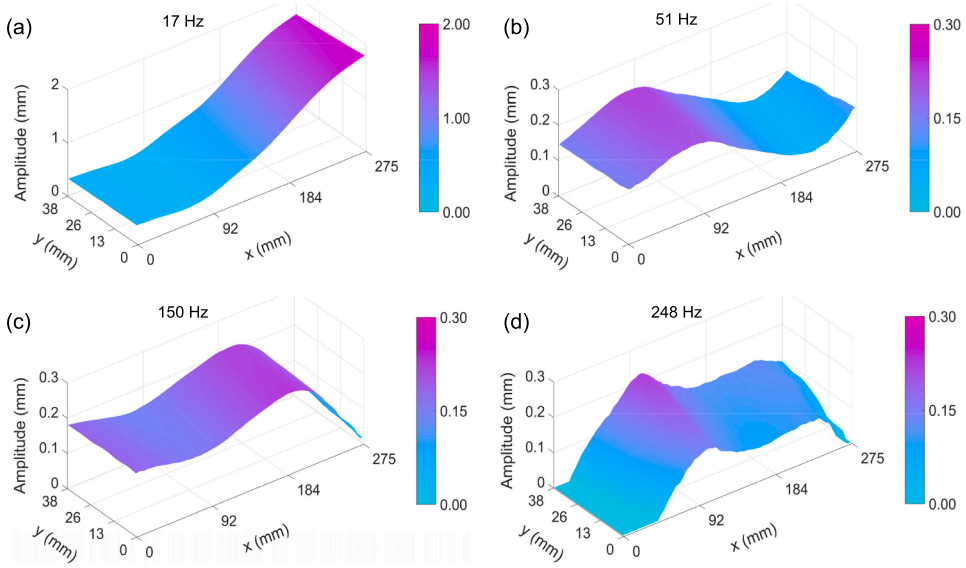

Fig. 9. Demonstration of the identified mode shapes. 3D mode shapes obtained at several sampled peak frequencies of (a) 17.0 Hz, (b) 51.2 Hz, (c) 150.6 Hz, and (d) 248.8 Hz.

tion (SVD), details of which can be found in [58,59]. The identified modal shapes are shown in Fig. 9. The results in this section suggest that estimating ODS fields with the proposed 3D optical sensing method is both feasible and accurate. In contrast to the traditional sensors (such as displacement transducers) which sense the dynamic response at single points, our method is also capable of identifying vibration frequencies and 3D modal shapes simultaneously.

Demonstration of damage identification. By applying the introduced wavelet transformation method of the ODS manifold in Section 2.3, we obtained the normalized damage response as shown in Fig. 10(a). An apparent peak value can be observed at the coordinates of $(170.13, 21.67)$ *mm* of the damage landscape relative to the predefined ROI, as shown in Fig. 10(b). To automatically identify damage, peaks and their locations can be detected by applying non-maximum suppression followed by peak detection. In practice, we expect to know the location of a defect on the structural component rather than its ROI-based coordinates. Hence, the damage coordinates in the ROI should be transformed to the object frame defined on the specimen. Let the lower-right corner of the specimen in Fig. 4 be the origin of the object frame, with the X- and Y-axis pointing upward and leftward from the origin, respectively, thereby defining the object frame relative to the world frame (i.e., $\mathbf{O}-XYZ$ in Fig. 2) according to the right-hand rule. In this object coordinate frame definition, we constructed the direction vectors of the X- and Y-axis by specifying the upper-right and lower-left corners of the specimen, denoted by $\mathbf{e}_x$ and $\mathbf{e}_y$, respectively. The direction of the Z-axis is thus given by $\mathbf{e}_z = \mathbf{e}_x \times \mathbf{e}_y$. The three corners defining the object frame are reconstructed by correspondingly selecting their projections in the left and right reference images and then performing standard triangulation. Denoting the origin of the object frame as $\mathbf{Q}$, the damage position in the object reference frame is given by

$$\mathbf{P}' = [\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z]^{-1}(\mathbf{P} - \mathbf{Q}), \tag{21}$$

where $\mathbf{P}$ is the reconstructed 3D position of the damage in the world frame. As a result, we obtain the final result as $(24.26, 195.13)$ *mm*, which is in agreement with the actual position $(25.00, 196.00)$ *mm* given in Section 3.1, and has acceptable precision, with a maximum positioning error of 0.87 *mm* in the vertical (length) direction. It should be noted that the object frame we define here is mainly to facilitate the verification of the accuracy of damage identification based on prefabricated geometric information. In practical applications, the appropriate object coordinate system is usually defined according to the needs of subsequent analysis (such as life assessment) to determine the damage location according to Eq. (21).

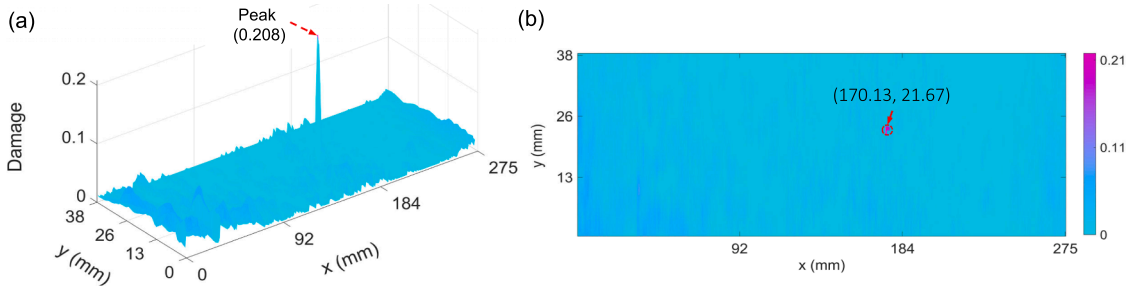

Fig. 10. Result of damage identification with the wavelet transform. (a) The identified damage landscape and its peak response, where 0.208 denote the degree of the identified damage; (b) the damage location $(170.13, 21.67)$ *mm* in the ROI-based coordinate frame.

For the degree of the identified damage, our algorithm determined it to be 20.8 %. To verify its correctness, we computed the degree of theoretical damage according to Eq. (19) using the stiffness parameter (detailed implementation can be found in [56] and elementary mechanics of materials), and obtained a value of 18.9 %. We can see that the degree of measured damage is in good agreement with the theoretical value, with an error of 1.9 %. These results demonstrate that the introduced damage identification algorithm in Section 2.3 is both feasible and accurate in identifying damage location and degree. Moreover, since the damage identification relies directly on the ODS estimates, this demonstration also provides comprehensive validation for our ODS estimation method as well as the calibration method.

3.5. Experimental demonstration on complex blade structure

In addition to the performance validation, we showcase the rigorous and broad applicability of the proposed geometric optical sensing framework by measuring the ODS and identifying the damages of a special-shaped blade structure, as shown in Fig. 11. The blade was made of white resin material (9400E) using 3D printing technology, with a fabrication tolerance of ± 0.1 *mm*, and has two prefabricated internal defects with diameters of 5 *mm* and 3 *mm* and depths of 0.8 *mm* and 0.6 *mm*, respectively, as shown by the red circles. It can be seen that the blade has a relatively complex irregular geometric structure (the chord length is 340 *mm*, the maximum and minimum widths are 120 *mm* and 60 *mm* respectively, and the maximum and minimum thicknesses are 6 *mm* and 1 *mm* respectively) and thus has 3D curved surfaces on both sides for allowing us to define a general ROI, as shown in the yellow area, for subsequent ODS computation. The experimental setup is the same as in Fig. 4 and the blade was fixedly mounted on the support through the root mounting hole. In the experiment, the blade was subjected to a random sinusoidal sweep excitation of 0–300 Hz applied by the exciter, which was positioned at the middle point 110 *mm* from the root of the blade, and the images of the blade motion were simultaneously captured

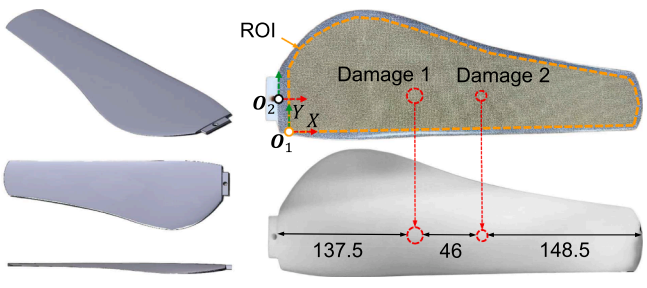

Fig. 11. Special-shaped blade specimen. On the left are three blade models from different viewpoints, and on the right is a surface images of the blade with the internal damage locations marked, where the top surface is patterned with a speckle pattern.The labeled distances to the embedded damages are measured in millimeters.

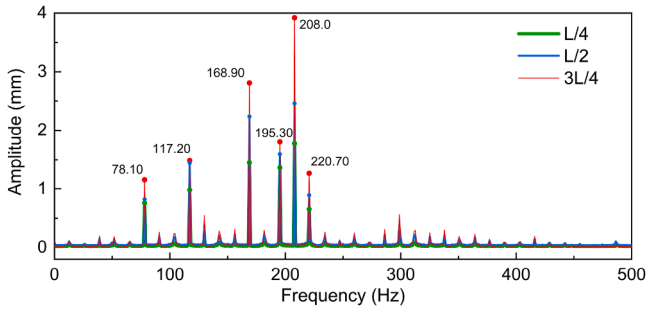

Fig. 12. Demonstration of identified ODS spectra. ODS spectra measured at three locations along the blade structure, specifically at one-quarter ($\frac{1}{4}$), one-half ($\frac{1}{2}$), and three-quarters ($\frac{3}{4}$) of the chord length L.

by the stereo camera for 2 seconds at the frame rate of 1000 fps after calibration.

By applying the proposed ODS sensing method in Section 2.2, we can obtain the ODS fields of the blade structure from the observed stereo image sequence. To observe the dynamic response characteristics of the blade structure under the aforementioned vibrational excitation, we selected sampling points at 1/4, 1/2, and 3/4 of the chord length, and plotted their ODS responses, resulting in the frequency spectrum curves shown in Fig. 12. We can observe that the frequency ranges where the ODS significantly enhance are identical at different spatial locations, therefore determining the six characteristic frequencies of the measured blade structure, as marked by the red dots, whose values in perfect

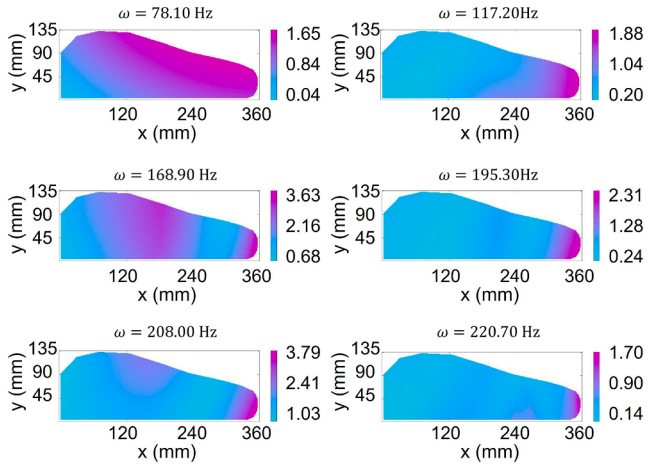

Fig. 13. Demonstration of the sensed ODS maps at six characteristic frequencies. Visualization of the full-field ODS maps captured at the six representative frequencies corresponding to Fig. 12, highlighting the spatial vibration patterns of the structure.

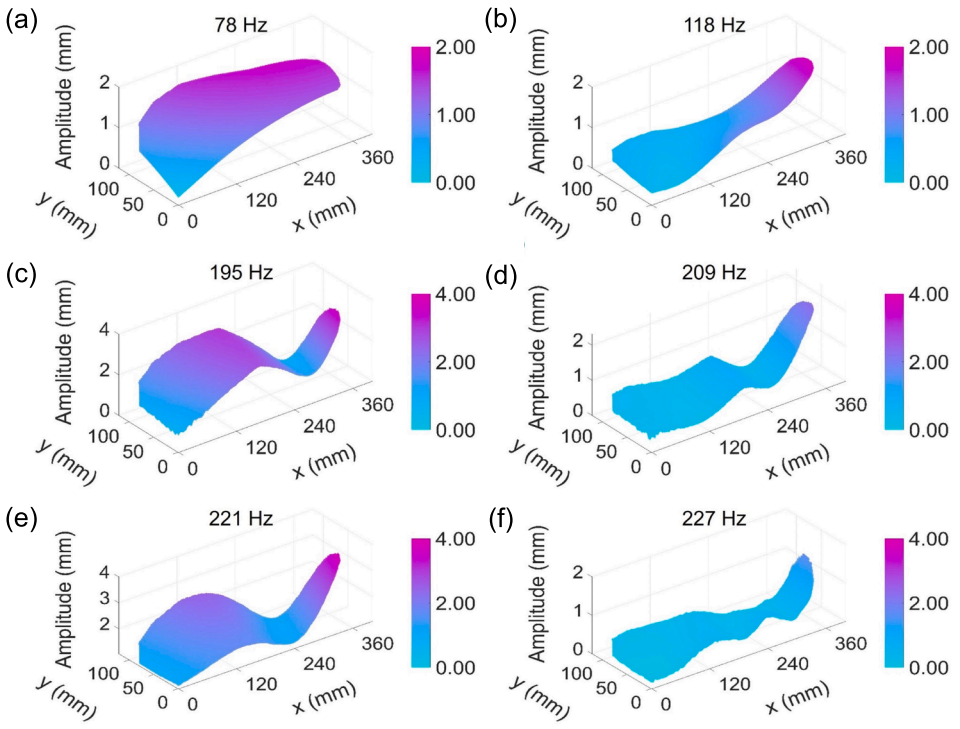

Fig. 14. Demonstration of identified mode shapes for the blade specimen. The mode shape corresponding to each labeled frequency is extracted from the respective ODS map presented in Fig. 13.

agreement with the theoretical results. To demonstrate the full-field dynamic deformation, we visualized the sensed ODS fields corresponding to these characteristic frequencies in Fig. 13 as well as the deduced 3D mode shapes from each of the ODS fields in Fig. 14. This result allows us to clearly observe the dynamic deformation of the blade structure from the 1st to the 6th order, as shown in Fig. 14 (a)-(f) respectively, without relying on specific excitation information.

Finally, we successfully identified the damages from the sensed ODS information according to the algorithm introduced in Section 2.3. The results are shown in Fig. 15. The identified degree values of damage 1 and 2 are 5.1 % and 3.2 %, respectively, which are consistent with the theoretical damage degree in terms of stiffness. Considering the irregularities of the blade geometry, we choose the lower-left corner of the ROI in Fig. 11 as the origin, denoted as $\mathbf{O}_1$, to construct a reference frame for determining the location of each damage, with the X- and Y-axis pointing leftward and upward from the origin, respectively. With this object frame configuration, the positions of the detected damage 1 and 2 can be transformed to the blade according to the method introduced in Section 3.4, which are $(127.08, 29.19)$ *mm* and $(180.07, 29.89)$ *mm*, respectively. To verify the correctness, we compared the localization results with the corresponding actual positions of the two damages relative to the reference frame with the origin at $\mathbf{O}_2$ in Fig. 11, which are $(126, 30)$ *mm* and $(179, 31)$ *mm*, respectively. It is found that the maximum positioning error of damage 1 occurs in the chord length direction, which is 1.08 *mm*, while the maximum positioning error of damage 2 occurs in the width direction, which is 1.11 *mm*. This shows that our damage identification method has the ability to accurately identify multiple damages. The experimental results demonstrate that the proposed geometric optical sensing framework offers potential portability and broader applicability, making it suitable for measuring ODS fields of and identifying damages in complex structural components even if subjected to randomly high-frequency vibrations.

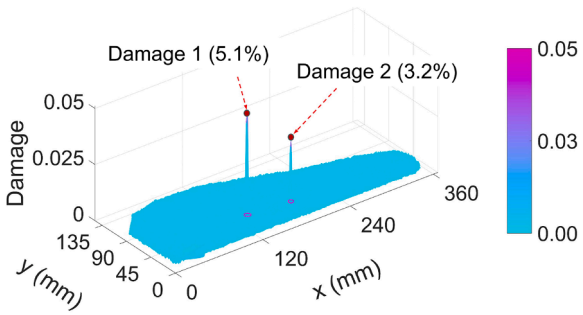

Fig. 15. Results of the identified damage locations and degree values. The identified damages, labeled as Damage 1 and Damage 2, are quantified by their relative stiffness loss as 5.1 % and 3.2 %. Their locations are marked by circles at coordinates $(127.08, 29.19)$ *mm* and $(180.07, 29.89)$ *mm*, respectively.

4. Discussion and limitations

In Section 3, we validate the feasibility and performance of the proposed method in a laboratory environment to provide a prototype of geometric optical ODS sensing framework with a fixed two-view geometry. From engineering perspective, we could develop an integrated portable stereo imaging device to apply the proposed method to practical engineering fields by integrating two compact high-speed CMOS sensors with an adjustable high-intensity light source via a biprism stereo optics [60], in which an embedded computing platform can also be integrated to ensure hardware-level synchronization and to run the proposed ODS sensing and damage identification algorithms. This integrated sensing system is feasible because the fixed stereo optics can be pre-calibrated once as described in Section 2.1 can then be used anywhere. In addition, our single-parameter ODS sensing algorithm has a low computational cost, making current image sensors and embedded

computing chips sufficient to support high-speed measurement tasks of several kilohertz.

However, for current version, the proposed method has two limiting factors that need further exploration to improve its implementation for long-term ODS and damage identification. The first factor is that the integration of the stereo imaging system used in this work is insufficient to handle long-term ODS monitoring tasks. The stereo vision system used in the experiments consists of two separate cameras. Directly applying this system to long-term ODS measurements presents challenges in system installation. Therefore, the compact biprism stereo optics as we mentioned before is a possible solution to forward it available for long-term monitoring as it has advantages including controllable illumination, high-precision synchronization, portability, and scalability to large-scale structures by using suitable lenses in a long working distance. The second factor concerns the damage identification algorithm. The damage identification algorithm in this work focuses on recognizing the location and extent of internal structural damage through detecting the impulse response signals encoded in the ODS manifold. Although experiments have validated that the method can simultaneously identify multiple damages, it cannot reflect geometric parameters such as the shape or orientation of the damage.

5. Conclusions

In this paper, we present a technical framework for full-field ODS sensing and damage identification in operating structures, leveraging a two-view stereo imaging system. The framework is approached upon several critical algorithms including RANSAC-based imaging system calibration, frequency-domain ODS estimation with depth parameterization, and damage identification based on the wavelet analysis of the ODS fields. To efficiently estimate the ODS fields, we approach a geometric vision sensing method that maps the motion of object points measured in stereo image space to the frequency domain by proposing a vector-based single-parameter linear time-frequency transform model.

We demonstrate the performance of the proposed algorithms through a series of experiments. The calibration accuracy of the stereo imaging system can be enhanced by the RANSAC-based calibration method, which guarantees the quality of the subsequent stereo rectification, while we have also demonstrated the accuracy of the proposed geometric optical sensing method by reconstructing the depth and disparity maps. Based on accuracy verification results, we showcase the ODS fields estimated by the proposed method and its capability in identifying frequency spectrum and modal shapes for structural components, without prior knowledge of the structural and excitation parameters. A comprehensive validation of our algorithms can be found in the final damage identification demonstration in which the damage location and degree are identified successfully for the tested plate and blade specimens.

Computer vision sensing techniques have already found potential applications in many fields involving full-field vibration measurement and damage identification, not least for those operating structures. With the proposed method, we can rapidly obtain the vibration responses, such as frequencies and modal shapes, and identify the damage parameters in industrial components and civil structures during operation. This may lead the method into even more impactful automation applications in structural health monitoring by incorporating some advanced stereo imaging modules and edge monitoring systems.

CRediT authorship contribution statement

Haoran Zhang: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation; **Jiaren Zuo:** Writing – original draft, Validation, Software, Methodology, Investigation; **Zhangyu Yao:** Visualization, Validation; **Deyang Zhang:** Resources, Funding acquisition; **Lei Lu:** Supervision, Resources, Funding acquisition; **Wei Pan:** Software, Investigation, Data curation; **Zhilong Su:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Data availability

Data will be made available on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (12572209), (12002197), (62375078), Shanghai Municipal Science and Technology Program Project (24ZR1423100), and the National Key Laboratory of Ship Structural Safety (NAKLAS2024KF016-S). Data and code will be made available in our Matrice repository on GitHub, and we recommend contacting the corresponding author for assistance.

References

[1] Y. Wen-quan, W. Zhi-hui, S. Yong-xin, Six degrees of freedom vibration transmissibility measurement of flexural opto-mechanical structures with 3-axis acceleration sensors, Precis. Eng. 38 (2014) 561–568.
[2] G. Ausanio, A.C. Barone, C. Hison, V. Iannotti, C. Luponio, L. Lanotte, Mechanical vibration sensor based on elastomagnetic composite, Sens. Actuators, A 129 (2006) 25–28.
[3] Z. Qiu, M. Mai, Y. Teng, X. Wang, J. Ma, Bearing-based FBG acceleration sensor for low-frequency vibration measurement, Measurement 228 (2024) 114322.
[4] F. Boden, B. Stasicki, M. Szypuła, P. Ružička, Z. Tvrdik, K. Ludwikowski, In-flight measurements of propeller blade deformation on a VUT100 cobra aeroplane using a co-rotating camera system, Meas. Sci. Technol. 27 (7) (2016) 074013.
[5] P.L. Reu, D.P. Rohe, L.D. Jacobs, Comparison of DIC and LDV for practical vibration and modal measurements, Mech. Syst. Signal Process. 86 (2017) 2–16.
[6] D. Gorjup, J. Slavič, M. Boltežar, Frequency domain triangulation for full-field 3D operating-deflection-shape identification, Mech. Syst. Signal Process. 133 (2019) 106287.
[7] Y. Hu, W. Guo, W. Zhu, Y. Xu, Local damage detection of membranes based on bayesian operational modal analysis and three-dimensional digital image correlation, Mech. Syst. Signal Process. 131 (2019) 633–648. https://doi.org/10.1016/j.ymssp.2019.04.051
[8] J. Huang, S. Paganoni, E. Zappa, 3D-DIC For large displacement with laser-Based speckle patterns, IEEE Trans. Instrum. Meas. 72 (2023) 1–8. https://doi.org/10.1109/TIM.2023.3307763
[9] Y. Xie, S. Xu, L. Yuan, M. Chen, P. Wang, J. Zhang, Dynamic responses of laminated and graded ZrC-Mo composites, Int. J. Mech. Sci. 271 (2024) 109134. https://doi.org/10.1016/j.ijmecsci.2024.109134
[10] D. Miller, K. Shah, S. Sockalingam, M.A. Sutton, F.D. Thomas, K. Kodagali, Uncertainty analysis of acceleration and velocity fields measured via DIC, Int. J. Mech. Sci. 302 (2025) 110599. https://doi.org/10.1016/j.ijmecsci.2025.110599
[11] A. Abdela, B. Buffel, B. Sirahbizu, F. Desplentere, Investigating the effects of alkalization on mechanical properties of enset fibers using digital image correlation technique, Result. Eng. 25 (2025) 103940. https://doi.org/10.1016/j.rineng.2025.103940
[12] T. Durand-Texte, M. Melon, E. Simonetto, S. Durand, M.-H. Moulet, Single-camera single-axis vision method applied to measure vibrations, J. Sound Vib. 465 (2020) 115012.
[13] K. Čufar, J. Slavič, M. Boltežar, Mode-shape magnification in high-speed camera measurements, Mech. Syst. Signal Process. 213 (2024) 111336. https://doi.org/10.1016/j.ymssp.2024.111336
[14] G. Lin, Z.-R. Lu, J. Liu, L. Wang, Monocular out-of-plane vibration measurement using parametric homography, J. Sound Vib. 570 (2024) 118140.
[15] J. Lee, K.-J. Park, Y.-H. Park, Enhanced video motion magnification utilizing envelop shift for vision-Based vibration measurements, IEEE Trans. Instrum. Meas. 73 (2024) 1–14. https://doi.org/10.1109/TIM.2024.3381257
[16] Q. He, S. Wang, Improving 2D displacement accuracy in bridge vibration measurement with color space fusion and super resolution, Adv. Eng. Inf. 65 (2025) 103248. https://doi.org/10.1016/j.aei.2025.103248
[17] D. Tan, J. Li, H. Hao, Z. Nie, Target-free vision-based approach for modal identification of a simply-supported bridge, Eng. Struct. 279 (2023) 115586. https://doi.org/10.1016/j.engstruct.2022.115586
[18] K. Feng, D. Hester, S. Taylor, C. O'Higgins, A. Ferguson, Z. Zhu, G. Zou, M. Lydon, J. Early, Experimental modal identification of a pedestrian bridge through drive-by monitoring integrated with shared-mobility vehicles, Dev. Built Environ. 20 (2024) 100562. https://doi.org/10.1016/j.dibe.2024.100562

[19] F. Peng, R. Zhao, K. Yu, G. Jiang, T. Li, S. Wang, J. Li, B. Ma, Output-only modal identification of full-field time-domain data for heated hybrid hollow sandwich structures, Int. J. Mech. Sci. 283 (2024) 109628. https://doi.org/10.1016/j.ijmecsci.2024.109628

[20] S.R. Varedi, B. Buffel, F. Desplentere, A digital twin framework for visco-Hyperelasticity calibration: experiment and simulation, Int. J. Mech. Sci. 296 (2025) 110310. https://doi.org/10.1016/j.ijmecsci.2025.110310

[21] W. Ji, K. Luo, K. Luo, Review on computer vision-based inspection and monitoring for bridge cables, Measurement 248 (2025) 116892. https://doi.org/10.1016/j.measurement.2025.116892

[22] Z. Su, J. Pan, L. Lu, M. Dai, X. He, D. Zhang, Refractive three-dimensional reconstruction for underwater stereo digital image correlation, Opt. Express 29 (8) (2021) 12131–12144.

[23] Z. Su, J. Pan, S. Zhang, S. Wu, Q. Yu, D. Zhang, Characterizing dynamic deformation of marine propeller blades with stroboscopic stereo digital image correlation, Mech. Syst. Signal Process. 162 (2022) 108072.

[24] K. Zaletelj, V. Agrež, J. Slavič, R. Petkovšek, M. Boltežar, Laser-light speckle formation for deflection-shape identification using digital image correlation, Mech. Syst. Signal Process. 161 (2021) 107899.

[25] R.I. Hartley, P. Sturm, Triangulation, Comput. Vis. Image Understand. 68 (2) (1997) 146–157. https://doi.org/10.1006/cviu.1997.0547

[26] Z. Peng, X. Wang, Z. Wang, W. Liu, M. Liu, Menglian, Evaluation of the impact of filter types and parameters upon the accuracy of phase-based optical flow method with a complex steerable pyramid, PLoS ONE 19 (9) (2024) 1–26. https://doi.org/10.1371/journal.pone.0308943

[27] M. Cao, Z. Su, T. Deng, W. Xu, Nonlinear pseudo-force in “breathing” delamination to generate harmonics: a mechanism and application study, Int. J. Mech. Sci. 192 (2021) 106124. https://doi.org/10.1016/j.ijmecsci.2020.106124

[28] K. Zaletelj, D. Gorjup, J. Slavič, M. Boltežar, Multi-level curvature-based parametrization and model updating using a 3D full-field response, Mech. Syst. Signal Process. 187 (2023) 109927.

[29] D. Li, Q. Xie, X. Gong, Z. Yu, J. Xu, Y. Sun, J. Wang, Automatic defect detection of metro tunnel surfaces using a vision-based inspection system, Adv. Eng. Inf. 47 (2021) 101206. https://doi.org/10.1016/j.aei.2020.101206

[30] X. Dong, C. Zhang, J. Wang, Y. Chen, D. Wang, Real-time detection of surface cracking defects for large-sized stamped parts, Comput. Ind. 159–160 (2024) 104105. https://doi.org/10.1016/j.compind.2024.104105

[31] A. Koko, D. Salim, N. Leung, N. Spetsieris, S. Smith, D. England, T. Sui, T. Fry, Exploring short crack behaviour and fracture transition in 5052 aluminium alloy, Result. Eng. 26 (2025) 105303. https://doi.org/10.1016/j.rineng.2025.105303

[32] W. Fan, P. Qiao, A 2-D continuous wavelet transform of mode shape data for damage detection of plate structures, Int. J. Solids Struct. 46 (25) (2009) 4379–4395.

[33] M. Makki Alamdari, J. Li, B. Samali, Damage identification using 2-D discrete wavelet transform on extended operational mode shapes, Arch. Civil Mech. Eng. 15 (3) (2015) 698–710.

[34] M. Abdulkareem, N. Bakhary, M. Vafaei, N.M. Noor, R.N. Mohamed, Application of two-dimensional wavelet transform to detect damage in steel plate structures, Measurement 146 (2019) 912–923.

[35] S. Cao, H. Nian, J. Yan, Z. Lu, C. Xu, Modal analysis and damage localization in plate-type structures via TDD and PE methods based on the data of an integrated highspeed camera system, Mech. Syst. Signal Process. 178 (2022) 109309. https://doi.org/10.1016/j.ymssp.2022.109309

[36] C. Li, J. Zhou, X. Wu, T. Liu, Phase-based video vibration measurement and fault feature extraction method for compound faults of rolling bearings, Adv. Eng. Inf. 62 (2024) 102897. https://doi.org/10.1016/j.aei.2024.102897

[37] D. Gorjup, J. Slavič, A. Babnik, M. Boltežar, Still-camera multiview spectral optical flow imaging for 3D operating-deflection-shape identification, Mech. Syst. Signal Process. 152 (2021) 107456.

[38] Z. Zhang, A flexible new technique for camera calibration, IEEE Trans. Pattern Anal. Mach. Intell. 22 (2000) 1330–1334. https://doi.org/10.1109/34.888718

[39] B. Huang, Y. Tang, S. Ozdemir, H. Ling, A fast and flexible projector-Camera calibration system, IEEE Trans. Autom. Sci. Eng. 18 (2021) 1049–1063.

[40] Z. Su, L. Lu, S. Dong, F. Yang, X. He, Auto-calibration and real-time external parameter correction for stereo digital image correlation, Opt. Lasers Eng. 121 (2019) 46–53.

[41] M. Róg, A. Rzonca, The impact of photo overlap, the number of control points and the method of camera calibration on the accuracy of 3D model reconstruction, Geomatic. Environ. Eng. 15 (2000) 67.

[42] Z. Liu, B. Guan, Y. Shang, Q. Yu, L. Kneip, Line-based 6-Dof object pose estimation and tracking with an event camera, IEEE Trans. Image Process. 1 (2024) 1–16.

[43] H. Ha, M. Perdoch, H. Alismail, I.S. Kweon, Y. Sheikh, Deltille grids for geometric camera calibration, in: 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 5354–5362. https://doi.org/10.1109/ICCV.2017.571

[44] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM 24 (6) (1981) 381–395. https://doi.org/10.1145/358669.358692

[45] A.J. Lacey, N. Pinitkarn, N.A. Thacker, An evaluation of the performance of RANSAC algorithms for stereo camera calibrarion, in: British Machine Vision Conference, 2000. https://api.semanticscholar.org/CorpusID:14038918.

[46] F. Zhou, Y. Cui, Y. Wang, L. Liu, H. Gao, Accurate and robust estimation of camera parameters using RANSAC, Opt. Lasers Eng. 51 (3) (2013) 197–212. https://doi.org/10.1016/j.optlaseng.2012.10.012

[47] Z. Su, L. Lu, S. Dong, F. Yang, X. He, Auto-calibration and real-time external parameter correction for stereo digital image correlation, Opt. Lasers Eng. 121 (2019) 46–53. https://doi.org/10.1016/j.optlaseng.2019.03.018

[48] G. Bradski, A. Kaehler, Learning OpenCV: Computer Vision with the OpenCV Library, O’Reilly Media, 2008. https://books.google.com/books?id=seAgiOfu2EIC.

[49] C. Li, C. Cai, A calibration and real-Time object matching method for heterogeneous multi-Camera system, IEEE Trans. Instrum. Meas. 72 (2023) 1–12. https://doi.org/10.1109/TIM.2023.3260263

[50] F. Zhong, C. Quan, Stereo-rectification and homography-transform-based stereo matching methods for stereo digital image correlation, Measurement 173 (2021) 108635. https://doi.org/10.1016/j.measurement.2020.108635

[51] N.A.J. Lieven, D.J. Ewins, N.M.M. Maia, J.M.M. Silva, Modal analysis identification techniques, Philos. Trans. R. Soc. Lond. Ser. A: Math. Phys. Eng. Sci. 359 (1778) (2001) 29–40. https://doi.org/10.1098/rsta.2000.0712

[52] M.H. Richardson, P.L. McHargue, Operating deflection shapes from time versus frequency domain measurements, in: Proceedings of the International Modal Analysis Conference, SEM Society for Experimental Mechanics INC, 1993, pp. 581.

[53] J.A. Gubner, W.-B. Chang, Wavelet transforms for discrete-time periodic signals, Signal Process. 42 (2) (1995) 167–180.

[54] S.G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Anal. Mach. Intell. 11 (7) (1989) 674–693. https://doi.org/10.1109/34.192463

[55] Mallat, A wavelet tour of signal processing, Elsevier, 2 edition, 1999.

[56] G.Z. Voyiadjis, P.I. Kattan, A new class of damage variables in continuum damage mechanics, J. Eng. Mater. Technol. 134 (2012) 021016. https://doi.org/10.1115/1.4004422

[57] Z. Zhang, A flexible new technique for camera calibration, IEEE Trans. Pattern Anal. Mach. Intell. 22 (11) (2000) 1330–1334. https://doi.org/10.1109/34.888718

[58] R. Brincker, L. Zhang, P. Andersen, Modal identification of output-only systems using frequency domain decomposition, Smart Mater. Struct. 10 (2001) 441–445. https://api.semanticscholar.org/CorpusID:14168704.

[59] M. Batel, Operational modal analysis - Another way of doing modal testing, Sound Vib. 36 (2002) 22–27. https://api.semanticscholar.org/CorpusID:44519239.

[60] D.H. Lee, I.S. Kweon, R. Cipolla, A biprism-stereo camera system, in: Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149), 1, 1999, p. 87 Vol. 1. https://doi.org/10.1109/CVPR.1999.786921