A multiple camera position approach for accurate displacement measurement using computer vision

Rolands Kromanis1 · Prakash Kripakaran2

Received: 26 June 2020 / Revised: 13 January 2021 / Accepted: 9 February 2021 / Published online: 1 March 2021 © The Author(s) 2021

Abstract

Engineers can today capture high-resolution video recordings of bridge movements during routine visual inspections using modern smartphones and compile a historical archive over time. However, the recordings are likely to be from cameras of different makes, placed at varying positions. Previous studies have not explored whether such recordings can support monitoring of bridge condition. This is the focus of this study. It evaluates the feasibility of an imaging approach for condition assessment that is independent of the camera positions used for individual recordings. The proposed approach relies on the premise that spatial relationships between multiple structural features remain the same even when images of the structure are taken from different angles or camera positions. It employs coordinate transformation techniques, which use the identified features, to compute structural displacements from images. The proposed approach is applied to a laboratory beam, subject to static loading under various damage scenarios and recorded using multiple cameras in a range of positions. Results show that the response computed from the recordings are accurate, with 5% discrepancy in computed displacements relative to the mean. The approach is also demonstrated on a full-scale pedestrian suspension bridge. Vertical bridge movements, induced by forced excitations, are collected with two smartphones and an action camera. Analysis of the images shows that the measurement discrepancy in computed displacements is 6%.

Keywords Computer vision · Image processing · Displacement · Signal processing · Condition assessment · Damage detection · Laboratory beam · Suspension bridge

1 Introduction

In the past decade, smartphones with their integrated sensor and software technologies have undergone tremendous enhancements. Smartphone applications for civil infrastructure monitoring have been widely researched [1, 2]. Literature also covers a range of case studies when smartphones have been employed for bridge SHM [3, 4]. The cameras supported in smartphones have improved at such a rapid pace that today’s smartphones have the capability to capture images comparable to professional cameras. For example, Samsung S20 can record ultra-high-definition 4 K (3840×2160 pixel) and 8 K (7680×4320 pixel) videos at 60 and 24 frames per second (fps), respectively, and high-definition (1280×720 pixel) videos at 960 fps. These features are sufficient to capture static and dynamic response of bridges [5, 6]. Consequently, smartphones can be deployed during periodic visual inspections, with minimum cost and effort, to collect objective video data that can complement the subjective information typically recorded by bridge engineers. For example, Zhao et al. [7] concluded that cable forces of cable-supported structures can be estimated equally accurately from videos of fixed and handheld smartphones after processing measurement. The collected data can be archived forming a historical record of bridge responses to loading, and when complemented with suitable data interpretation tools, engineers can use the collected data to track and detect changes in structural conditions.

There are three broad steps in processing videos for structural displacement [8]—camera calibration, target tracking and displacement calculation. Many studies have developed image processing algorithms for measuring the motion of a single target on the structure and capturing structure’s static
effects can, however, be neglected if the measurements dur-
much larger than its response to traffic loads [26]. Thermal
bridge response to seasonal changes in temperature may be
dynamic response can mask early signs of damage [25]. Also
shown that the influence of temperature variations on bridge
collected at discrete time instants. Previous studies have
response can be an important factor in measurements
This paper will investigate and address this. Thermal effects
be compromised when images are collected in this manner.

gated whether the accuracy of structural response data may

tuations to a common coordinate system for bridge condition
Previous studies have not examined the feasibility of

damage detection. Particularly, studies have not investi-
gated whether the accuracy of structural response data may
be compromised when images are collected in this manner.
This paper will investigate and address this. Thermal effects
on response can be an important factor in measurements
collected at discrete time instants. Previous studies have
shown that the influence of temperature variations on bridge
dynamic response can mask early signs of damage [25]. Also
bridge response to seasonal changes in temperature may be
much larger than its response to traffic loads [26]. Thermal
effects can, however, be neglected if the measurements
during individual campaigns (i.e. inspections) are collected over
a short duration [27, 28] and if the emphasis of the data
interpretation is on the immediate response to static loading
rather than the quasi-static and dynamic response, as is the
case in this study. This is supported by previous studies that
have demonstrated the use of static load tests to assess the
condition of the structure and obtain bridge load ratings. For
example, Klaiber et al. [29] employed trucks to load a con-
crete girder bridge before and after damages were repaired.
Vertical deflections of girders were observed to reduce by
almost 20% after repair. Also, Dong et al. [17] used portable
cameras and computer vision technology to perform bridge
load rating.

This study will evaluate thus a vision-based approach that
analyses image recordings from cameras in different posi-
tions to accurately compute structural displacements. The
premise of this study is that the spatial relationships between
multiple structural features such as bolts in cast iron bridges
remain the same even when images of the structure are taken
from different angles or camera positions. This premise is
true as long as these features are located on the same struc-
tural plane. The study employs smartphone technologies to
investigate (i) if structural features can be accurately located
from images collected from different camera positions, and
(ii) if structural response can be accurately estimated. A tim-
ber beam with artificial structural features served as a test-
bed. While the beam is undergoing load tests, smartphones
are used to collected images, from which target locations are
obtained and transformed to the structural reference plane.
The proposed measurement collection approach is also vali-
dated on a full-scale pedestrian bridge subjected to forced
excitations.

2 Methodology

An approach to vision-based deformation monitoring that
is independent of camera positions used to collect the data
for the condition assessment of bridges is developed in this
paper. Figure 1 illustrates the proposed approach, which
includes the following steps:

1. Image collection and processing
2. Structural response generation
3. Structure’s condition assessment

The initial set of collected responses may be taken as
representative of the structure’s baseline (normal) condi-
tions. A change in structure’s conditions can be detected
from new measurements by comparing these against the
baseline response. This can enable the asset owner to plan
an intervention such as a detailed inspection to ascertain the
underlying reason for the change. The following sections
describe the above-mentioned steps in further detail.

2.1 Image collection and processing

Successful image processing relies on multiple factors, one
of them being reliable data collection. In short-term (i.e.,
lasting from a few seconds to a few minutes) measurement
collection events, camera stability and accurate camera focus
are easily ensured. However, positioning cameras to ensure
exactly the same field of view during all measurement col-
lection events, which may be separated by months, is very
difficult. Consequently, to ensure measurements taken at
different events are comparable, all data need to be trans-
formed to the same coordinate system. This can be done
in two stages: (i) generate a planar homography matrix
to transform coordinates in the current bridge coordinate
plane (i.e., as captured in the image) to a defined reference
plane (i.e., as provided in structural plans) and (ii) apply the
matrix to convert target locations in the collected images to
the defined reference plane. The planar homography can be
applied when targets on the structure move within a single
plane. The projection relationship between two dimensional
(2D) structural and image planes is given in Eq. 1 [8].

\[ \alpha \{u\} = [P]_{3 \times 3} \{X_P\}, \]

where \(X_P\) is a 2D structural plane \((X_P = [X, Y, 1]^T)\), \(u\) is a
2D image plane \((u = [u, v, 1]^T)\), \(P\) is the planar homography
matrix and \(\alpha\) is an arbitrary coefficient. The two stages (and
their respective steps) are described below in further detail.

Stage 1: Generation of the planar homography

1. Define reference points These points, also called control
points, are essentially visual patterns that can be clearly
discerned from images of the structure. The coordinates
for these points are defined from a 2D structural drawing
of the bridge or by physical measurement in the field. As
a minimum four reference points are needed for planar
homography, in which 2D points in the structural surface
plane are mapped to their corresponding points in the
image plane [9, 18, 30].

2. Generate a transformation matrix The reference points
defined in the previous step are located in the reference
image collected during the monitoring event. This may
be done using feature detection algorithms as outlined
in the next stage. The reference points and image points
can be input to a geometric transformation algorithm
such as fitgeotrans in MATLAB [31] to compute the
geometric transformation matrix.

Stage 2: Computation of structural response at target
locations

1. Select targets Targets can be selected either manually
or using an automated feature finding algorithm. Those
with multiple and distinctive features (e.g., corners,
edges, surface patterns) can be located easier, quicker
and with higher accuracy than targets that are blurry and
have unclear boundaries.

2. Derive target features An appropriate feature detection
algorithm such as Harris–Stephens algorithm [32] or
speeded-up robust feature detection [33] algorithm is
selected. Detected features (interest points) are assigned
to the targets that they describe and are sought in con-
secutive image frames.

3. Specify region of interest (ROI) Tracking targets with
small, predictable movements (such as in deformation
monitoring of bridges) can be much easier than other
tracking tasks (e.g. moving people) if a ROI, within
which a target is expected to move, is specified. This
can reduce computational time and avoid errors arising
from similar targets being present in the same image frame.

4. **Track and record coordinates of the selected targets** The coordinates of the targets in each image frame are identified. The target coordinates, which are indicative of target movements, along with the time and target number are stored in an array for computing structural response.

5. **Transform target locations to reference plane** The geometric transformation matrix (generated in Stage 1) is multiplied by image coordinates of targets to evaluate the coordinates for the targets in the reference coordinate plane. This step is repeated for all image frames. The coordinates can then be used to compute structural movements.

The first stage could also include an optimization step, in which (i) multiple targets, and their numbers and combinations, and/or (ii) transformation algorithms (e.g., perspective, affine, polynomial) are chosen for the matrix generation. The matrix accuracy can be evaluated using targets that are found on the structure but not included in the generation of the matrix. The main drawback of the optimization step is that increasing the number of reference targets and using complex algorithms can result in overtraining, i.e., generation of a matrix that works well only for the reference image in which the structure is not subject to loading [34]. However, when a load is applied and image reference points change locations, the locations of targets may become erroneous due to the overtraining of the transformation matrix.

### 2.2 Response generation

Movement of a target in the image frames is referred to as a target displacement. This displacement is a projection of the real in-situ movement onto the $x$–$y$ plane of the defined (reference) coordinate system. Note that a target may move along the longitudinal axis of the bridge as well as the vertical direction. For this reason, target movements are referred to as target displacements rather than deflections, which correspond to the vertical deflection of a bridge along its length. The vertical deflection ($\delta_v$) of the structure at a specific location can be calculated from the change in its $y$ coordinates as follows.

$$\delta_v = T_0[y] - T_n[y],$$  \hspace{1cm} (2)

where $T_0$ and $T_n$ are target coordinates before load is applied and at $n^{th}$ measurement, respectively.

Consecutively collected target displacements or deflections form response time histories or signals. Signals may be noisy (e.g., change in light conditions) or contain outliers (e.g., a moving obstacle in ROI at the image capture) thus, requiring applications of signal pre-processing techniques such as denoising (e.g., with moving average filter) and outlier removal (e.g., inter-quartile range analysis). Another signal pre-processing option is the selection of a known stable location in an image frame to correct camera movements [13].

### 2.3 Measurement accuracy and structural condition assessment

Measurement residual ($e$) for a response parameter ($r$) such as vertical deflection at $i^{th}$ and $j^{th}$ measurement events is expressed in Eq. 3. The measurement events can be referred to both response collection from the same or different camera(s) and camera position(s).

$$e = \frac{r_i - r_j}{r_i}.$$  \hspace{1cm} (3)

Measurement residuals can be used to assess (i) the accuracy of the structural response generated at different camera positions or measurement discrepancy/deviation (ii) and changes in conditions of the structure. The threshold for measurement residuals can be case specific and based on the judgement of an engineer. For example, $e \gg \pm 5\%$ may indicate that the condition of the structure has changed sufficiently to warrant further measurement analyses or inspections. Similarly, distinct troughs (or drops) in $e$ values of targets along the length of a bridge can be indicators of damage locations.

### 3 Laboratory experiments

In this section, the performance of the proposed vision-based approach is evaluated on a laboratory beam.

#### 3.1 Laboratory test setup

A simply supported timber beam subjected to static loads serves as a testbed. The beam is 1100 mm long, 25 mm wide and 45 mm deep (see Fig. 2). 43 artificial targets ($Ti$, $i = 1, \ldots, 43$), in a form of full circles, are drawn on the surface of the beam following a template shown in Fig. 2b. Only names of a few representative targets are provided in Fig. 2c. Targets are named sequentially from left to right, starting from the top left target. Previous studies by Kromanis et al. [35] demonstrate that structural deformations computed from target displacements deviate by utmost 2.5% from those evaluated using contact sensors. Therefore, in this study, beam deformations are captured only with smartphones.

#### 3.2 Measurement scenarios

The experimental procedure consists of manual application and removal of a load (100 N) at the centre of the beam
in the absence of and presence of damage. In experimental studies, beam structures have been damaged with section cuts up to 62% of the section area at cut locations [36–40]. In this study, there are three 20 mm deep and 45 mm long section cuts at the top (compression) side of the beam simulating damage. Tight-fit wooden blocks (Bs) fill the section cuts. The blocks are used to reassure the repeatability of the damage scenarios for multiple events at multiple camera positions. The blocks can be removed without the beam being disturbed. Although the blocks fit tightly the beam is not expected to perform as a solid beam, i.e., with no cut-outs. When all blocks are in place, the beam is healthy (no damage) and corresponding measurements represent baseline conditions. When a block is removed, a damage is created. The damage severity is regulated by the number of removed blocks. Measurements are taken for 20 s after loading or un-loading to allow for any vibrations to damp out. Smartphones are employed to capture images of the laboratory setup at 1 Hz. They are fixed on sturdy tripods, which can be assumed to be perfectly still.

Two scenarios are considered:

- **Scenario 1** This consists of a single damage event being measured using a number of cameras set up in different positions. The sequence of events involves loading of the undamaged beam, unloading, introduction of damage (by removal of a wooden block) within the beam and loading to measure deformations.

- **Scenario 2** This is similar to Scenario 1 but with multiple damage events that gradually increase the level of damage (by removing multiple wooden blocks). In this scenario, each sequence of events is measured using only one camera but with the camera position changing between the events.

The laboratory setup, image processing steps, response generation and damage detection for both scenarios are described in below sections.

### 3.3 Single event—multiple camera positions

The proposed monitoring approach is initially evaluated on a single event during which three smartphones collect beam deformations. Smartphone makes and camera specifications are given in Table 1. Smartphones are placed at different angles, heights and distances to the beam (see Fig. 3). The distance of each smartphone camera to T13 ($d_{T13}$), T28 ($d_{T28}$) and T36 ($d_{T36}$), and camera plan ($\alpha$) and side ($\beta$) view angles to T36 are given in Table 2. Negative $\alpha$ and $\beta$ indicate that the camera is positioned right and above T36, respectively. Figure 4 illustrates the distances and angles of the camera to the structure (beam). Images of the beam are collected at no load and 100 N load, both before and after damage, which is introduced by removing B2 (see Fig. 2).

#### 3.3.1 Image collection and processing

A semi-supervised image processing process is adopted to analyse collected images and calculate target displacements following the three stages in Sect. 2.1.

### Table 1 Smartphone makes and camera specifications

| Name | Make     | Camera and lens specifications         |
|------|----------|----------------------------------------|
| S1   | Samsung A3 | 8 MP, f/2.4 aperture, 31 mm (wide)    |
| S2   | Samsung A5 | 13 MP, f/2.0 aperture, 28 mm (wide)   |
| S3   | Samsung S8 | 12 MP, f/1.7 aperture, 26 mm (wide)   |
Generation of the geometric transformation matrix

Target locations are known. The horizontal distance from T13 to the left support is 100 mm (as measured and shown in Fig. 2c). The vertical distance from T29, T30,…, T43 to the bottom of the beam is 10 mm. Four reference points on x–y reference plane, which correspond to T1, T12, T29 and T43, are selected for the generation of the geometric projection matrix. Target locations on the image plane are calculated using image processing analysis in the previous stage. The projective transformation has been shown to generate accurate planar homography matrices in previous studies [30, 34, 41] and, therefore, it is selected.

Computation of target locations

Figure 5 illustrates the steps using a cropped region of the first image captured with S1 as an example. The steps are discussed below.

(a) The Hough transform method for finding circles is suitable for the target detection and location [42]. A search region, which is the beam surface facing the camera, is specified to optimize the search area and time, and reduce the number of circles found in the image. The range of search radii is the main criteria in the Hough transform method. It is expressed in the number of pixels. The farther a camera is positioned from the centre of the beam the larger must be the range of radii. The circle detection method sorts targets based on parameters such as detection sensitivity and circle size. Figure 5a shows that the order of the detected circles is random and does not follow the numbering sequence defined in Fig. 2c. A sorting algorithm is used to arrange circles in rows and then columns to ensure that the identified targets are numbered as in Fig. 2c.

(b) The size of ROI assigned to a target is derived from the size of the detected radius of the circle and its position along the length of the beam. The targets located closer to the supports are expected to move less in the vertical direction during the application of the load than the targets closer to the centre of the beam. The range of a target displacement is found by comparing the first image, in which no load is applied to the beam, to an

![Table 2 Camera distances and angles relative to the beam](image)

| Smartphone | $d_{T13}$ (mm) | $d_{T36}$ (mm) | $d_{T28}$ (mm) | $\alpha$ (°) | $\beta$ (°) |
|------------|----------------|----------------|----------------|-------------|-------------|
| S1         | 1095           | 1000           | 1095           | 0.4         | 0           |
| S2         | 965            | 1105           | 1385           | 5.6         | 30.1        |
| S3         | 1365           | 1115           | 1025           | 10.3        | −24.2       |

![Fig. 4 The position of the camera relative to the structure](image)
image, in which the maximum vertical displacement is expected. In this scenario, maximum displacements of targets are when the load is applied to the damaged beam. Figure 5b illustrates ROIs, which are numbered according to the target numbers (see Fig. 2c), for targets located on the left side of the beam.

(c) The DeforMonit application technique developed at Nottingham Trent University by R Kromanis [43] is then used to evaluate target displacements. The technique is demonstrated in Fig. 5c, where (i) shows the original image as obtained from ROI, (ii) shows a grayscale image, in which the number of pixels is increased by a factor of four and the sharpness and contrast are adjusted, (iii) is a binary image, in which regions of pixels form blobs, and (iv) shows only the target (other blobs are removed) with an ellipse drawn around its boundary. The centre of the ellipse represents the centre of the target, which is recorded and passed to the next image processing stage.

**Transformation of target locations** The computed geometric transformation matrix is applied to all target locations (centres of blobs) found in the first image processing stage. An example of target transformation from S2 is shown in

---

**Fig. 5** Image processing steps: a automatic detection of targets, b setting regions of interest, c detection of target location in ROI1: (i) before processing, (ii) adjusted image, (iii) binary image, and (iv) binary image with the identified target (green ellipse) and its location (green ‘+’)

---

**Fig. 6** Targets for S2 before (top) and after (bottom) the projective transformation
Fig. 6. Although a slight ‘bow effect’ [blue line in Fig. 6 (bottom)] is discernible for the targets closer to the middle of the beam, this can be neglected since target displacements range only between a few millimetres/pixels. The camera intrinsic and lens distortion parameters were deliberately not considered for the camera calibration. Calibration in this study relies solely on the generation of the geometric transformation matrix for the following reasons: (1) different cameras might be employed during bridge monitoring events by inspectors, thus requiring a simple-to-use and robust approach, and (2) studies have shown that there is a negligible difference (< 0.6% of the vertical displacement range) between results obtained from raw and undistorted smartphone images [14, 35].

3.3.2 Response generation

Target displacements are converted to vertical deflections ($\delta_v$) and used as a damage sensitive structural response parameter. Raw and pre-processed vertical deflections at T36 are shown in Fig. 7. Noisy deflections with an upward drift are observed from images collected with S3. This may be due to the specific smartphone or its make. Deflections collected with the other two smartphones are less noisy and do not drift. The drift from S3 deflections can be removed either using a stationary reference target in the background or signal processing techniques. A signal processing technique, in which a 2nd order polynomial curve generated for the no-load period, is selected to remove the measurement drift. A moving averaging filter of 6 measurements is applied to deflections. Final (processed) vertical deflections at T36 derived from all smartphones are similar. The beam deflection continues to increase marginally with the presence of the load. For each target, a single deflection value, which is the average value between load application and removal [see amber shaded periods in Fig. 7 (right)], is taken forward to the condition assessment stage.

The beam deflection at 100 N load for all target locations is shown in Fig. 8 (left). The deformed shape represents the anticipated beam deflection. There is a target missing at the centre of the beam in the top row (T1–T12), where B2 is located; this is reflected in the plot. Taking the bottom row of targets (T29–T43) as representative of the beam deflection curve, Fig. 8 (right) plots the vertical deflections at the targets for all camera positions for undamaged and damaged beam. Deflection curves differ slightly for camera positions. The measurement residual for camera positions and damage detection are analysed in the next section.

3.3.3 Condition assessment

The beam is grade C16 timber, which has a mean elastic modulus ($E$) of 8.0 kN/mm$^2$ [44]. The maximum deflection is at mid-span when the point load ($P$) is applied at the middle of the simply supported beam. Assuming linear elastic behaviour and small deformations, the mid-span deflection can be calculated using Eq. 4.

$$\delta = \frac{P l^3}{48 E I}.$$  

where $l$ is the length of the beam and $I$ is the second moment of area. Rearranging terms in the equation, the overall $E$ for the experimental beam at a healthy state is 4.5 kN/mm$^2$, which is 44% smaller than the given value. This indicates that the beam at its healthy state, when the cut-out blocks are in place, already does not perform as a solid timber beam.

The condition of the beam is analysed using response (vertical deflection) measurements computed in the response generation step. The residual $e_{\delta_v}$ between vertical deflections...
computed from \( i \)th and \( j \)th cameras at a selected target is derived as follows:

\[
\epsilon_{\delta_{y}} = \frac{\delta_{V,i} - \delta_{V,j}}{\delta_{V,i}}.
\]  

(5)

\( \epsilon_{\delta_{y}} \) values for targets distributed along the length of the beam can be graphically plotted. For example, \( \epsilon_{S1S2} \) represents the line of residuals for cameras S1 and S2. In the plots of \( \epsilon_{S1S2} \) and \( \epsilon_{S1S3} \) at no damage (Fig. 9a) and damage (Fig. 9b) states, the assumption is that measurements from S1 represent beam baseline conditions. At no damage state, \( \epsilon_{\delta_{y}} \) values do not exceed \( \pm 5\% \) confirming that the image processing and response generation steps provide an accurate structural response in multiple camera positions. For targets with small vertical displacements (i.e., targets near the supports) \( \epsilon_{\delta_{y}} \) values are higher. At damage state, all \( \epsilon_{\delta_{y}} \) values surpass \( -20\% \) and \( \epsilon_{\delta_{y}} \) through \( -34\% \) is at the damage location (midspan of the beam). Figure 9c plots \( \epsilon_{S1} \), \( \epsilon_{S2} \) and \( \epsilon_{S3} \), which are measurement residuals computed between deflection measurements from the same camera for the undamaged and damaged beam with loading. Measurement residuals drop at the mid-span of the beam and the overall results demonstrate the reliability of using measurement residual as a damage sensitive parameter for damage detection and location.

Fig. 8 Vertical deflections of the beam captured with S1 at different heights (left) and S1–S3 (left) at no damage (solid lines) and damage (dashed lines)

Fig. 9 Measurement error (\( \epsilon_{\delta_{y}} \)): a no damage, b damage, c for all smartphones individually undamaged–damaged
3.4 Multiple events—multiple camera positions

This section provides results of multiple events captured at six camera positions ($P_i$, $i = 1, \ldots, 6$). A single camera (Samsung S5) is used. A ghost image showing all the camera positions and their corresponding views is given in Fig. 10. Table 3 provides camera distance to the three targets on the beam and two camera angles with respect to $T_{36}$. At $P_1$ the beam represents baseline conditions. The other camera positions capture the following damage scenarios: $D_1$ (B1 removed), $D_2$ (B1 and B2 removed) and $D_3$ (B1, B2 and B3 removed). For brevity, this section omits the image processing and response generation steps, which are the same as those described and demonstrated in Sect. 3.3.

3.4.1 Condition assessment

Vertical deflections at the mid-span of the beam measured at $T_{36}$ during experimental testing for all six camera positions are shown in Fig. 11. Discrepancies are observed in deflections for periods when load is applied at both healthy and damaged state of the beam. Due to the nature of the experimental setup, the duration of load application for some events is shorter/longer than for other events. A visible change in deflections is observed between no damage and $D_2$. Between no damage and $D_1$, and $D_2$ and $D_3$ the deflection difference is small requiring a closer assessment.

Vertical deflections along the length of the beam for all camera positions at no damage and $D_1$ are given in Fig. 12 (left). Although the figure is saturated with beam deflections from all camera positions for two scenarios, a discernible change can be observed in beam deflections for $D_1$ (dashed lines). It is also noticeable that deflections at $P_6$ for no damage scenario are even larger than deflections measured at different locations for $D_1$, especially for the right side of the beam. Figure 12 (right) plots $e_{P_i}$ for all positions using $P_1$ as the baseline. The plot shows that measurements from $P_6$ ($e_{P_1P_6}$) deviate the most and $e_{P_1P_2}$ has the smallest measurement residual.
Measurement residuals for target deflections computed using P1 and \( j \)th camera positions for various damage scenarios are plotted in Fig. 13. Although the measurement residuals between camera positions (from \( e_{P1P4} \) to \( e_{P1P6} \)) show a degree of variation for a specific scenario, the extent of change is much larger for a damage scenario than for an undamaged scenario. Maximum \( e_\delta \) values are observed to be in the region of damage. For example, at D1 a damage

![Graph of measurement residuals for target deflections computed using P1 and \( j \)th camera positions for various damage scenarios.](image)

**Fig. 12** Left: vertical deflections of the beam at no damage (solid lines) and D1 (dashed lines). Right: measurement error for the ‘no damage’ scenario

![Graph of measurement error along the length of the beam: a D1, b D2, and c D3.](image)

**Fig. 13** Measurement error along the length of the beam: a D1, b D2, and c D3

![Graph of measurement error statistics for all scenarios.](image)

**Fig. 14** Measurement error statistics for all scenarios (color figure online)
is created close to the left side support. Peak measurement residuals in Fig. 13a plots are correspondingly concentrated on the left side. When the damage is created at the mid-span of the beam (i.e., D2) and right to the mid-span measurement residual throughs shift.

Root-mean square deviation (RMSD) is derived from measurement residuals between P1 and jth position for a number of targets (n) along the bottom of the beam using Eq. 6: RMSD gives an overview of the overall measurement residual or discrepancy in computed displacements relative to the mean. Figure 14 provides a bar plot of RMSD of measurement residuals, together with the maximum residual and their locations for a corresponding camera position. Only RMSD of eP1P0 exceeds 5% threshold (a thick red line in Fig. 14) at no damage scenario. The measurement residual and RMSD residual analysis in Fig. 14, show that measurements from P6 are erroneous and exceed the damage threshold even when the structure is not damaged. This could be related to the camera angle θ, which is almost three times larger than that for other positions.

\[
\text{RMSD} = \sqrt{\frac{\sum_{i=1}^{n} (e_{P1,k} - e_{j,k})^2}{n}}. \tag{6}
\]

Damage can be accurately located, when analysing measurement residuals for each camera position separately between two consecutive events. Figure 15 plots these measurement residuals between (i) no damage and D1, (ii) D1 and D2, (iii) and D2 and D3 scenarios. The average residual for each scenario is superimposed on the residuals for each camera position with thick lines. The damage is located in the position where residuals are the lowest, for example, for No damage—D1 combination the residual drops at around a 250 mm mark. Damage locations are shown in Fig. 2a.

4 Wilford Suspension Bridge monitoring

The accuracy of the proposed measurement collection approach is investigated on the Wilford Suspension Bridge. The bridge spans 69 m linking Nottingham to West Bridgford over the River Trent. It is both a pedestrian bridge and a water aqueduct. The bridge is subjected to a range of 60 s long forced excitations (i.e., students jumping on the deck). The experiment is organized by the University of Nottingham as part of a student assignment. In this study, a scenario when students are jumping on the side of the deck that is closer to the camera positions is considered. Two smartphones (Samsung S8 (S1) and Samsung S9 (S2)) with 12 MP camera, and f/1.5–2.4 aperture and 26 mm (wide) lens) and a modified GoPro (GP) Hero 5 action camera with a varifocal zoom lens (25–135 mm) are positioned on the left river bank. All cameras record 4 K videos at 30 fps. Figure 16 (top) shows the monitoring set-up. The region of interest (ROI) contains approximately 14 m of the deck length at the middle of the bridge. The ROI has the same size as the frame of the GP. Seven targets (Ti,  i = 1, ..., 7) are selected in the ROI. These are shown in a sketch of the bridge in Fig. 16 (bottom). Distances of the cameras to T1, T4 and T7 together with camera angles to T4 (as shown in Fig. 10) are estimated and listed in Table 4.

Frames from S1, S2 and GP, and ROIs with the targets are shown in Fig. 17. Harris method [32] for detecting corner features is employed to detect features characterizing targets. The camera motion is removed using the displacements of a stationary target in the background [41]. The four reference points for the generation of the planar homography matrix correspond to the top and bottom ends of the balusters at T1 and T7 locations. The coordinates of the reference points are obtained from structural drawings of the bridge. Then pixel displacements are converted to structural displacements. Vertical displacements are pre-processed with a 5-s moving average filter, removing remaining camera movements. A 2 Hz high-pass filter is applied to remove the high-frequency noise. Measurement histories are manually synchronized giving a set start time. Vertical deflection time histories from all cameras for the entire excitation period and 21 s are shown in Fig. 18. The bridge first vertical mode at the studied excitation is at 1.63 Hz, which is the same as computed from measurements with Global Navigation Satellite System (GNSS) employed during the experiment.

The measurement accuracy is evaluated with RMSD of vertical deflections (Δvi) for each target computed between two cameras k and l using Eq. 7. In Eq. 8 measurement residuals (e) for each target is derived from the sum of RMSE
values and range of vertical (or peak-to-peak) displacements \((r_{V}/u1D6FFV)/u1D6FF\) from the three cameras \((k = 1, 2, 3)\). RMSE together with \(e\) values are plotted in Fig. 19. The measurement discrepancy in computed displacements is 5.9%, which is 0.9% higher than the set ± 5% damage indicating threshold in the

| Camera | \(d_{T1}\) (m) | \(d_{T4}\) (m) | \(d_{T7}\) (m) | \(\alpha\) (°) | \(\beta\) (°) |
|--------|----------------|----------------|----------------|----------------|----------------|
| S1     | 32.90          | 39.60          | 46.30          | 6.8            | 78.3           |
| S2     | 38.05          | 44.40          | 50.90          | 7.1            | 70.1           |
| GP     | 38.85          | 45.25          | 51.80          | 5.7            | 70.6           |

Fig. 16 The Wilford Suspension Bridge experimental setup (top) and a sketch of the bridge’s elevation (bottom)

Fig. 17 Camera frames and regions of interest. Yellow ‘x’ are reference points for the matrix transformation
laboratory studies, therefore suggesting that the threshold may need adjusting according to in-situ measurements.

$$\text{RMSD}_{k,l} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\delta_{V,k,i} - \delta_{V,l,i})^2},$$  \hspace{1cm} (7)

$$e = \frac{\sum \text{RMSD}_{k}}{\sum \delta_{V,k}}$$  \hspace{1cm} (8)

A single period of the bridge vertical motion (from 8.4 to 9 s) is analysed further to demonstrate the accuracy of the vision measurement and its relevance to the bridge condition assessment within the proposed approach. Figure 20 shows the vertical displacements of all target in the ROI of GP for the selected period. The range of displacements for each target is related to their position on the bridge. The target at the midspan of the bridge (T4) has the largest range. The range of vertical displacements reduces targets away from the mid-span of the bridge. The range of displacements of each target along the length of the bridge is given in Fig. 21. GP measurements are according to expected deflections of the superstructure, considering its geometry. The ranges of target displacements computed from S1 and S2 do not follow as accurately the anticipated deflection pattern as those from GP. Setting the GP measurements as the reference, the largest deviation for S1 and S2 are 5.4% and 9.1%, respectively, and for both cameras it is for T6. The relative mean deviation of S1 and S2 is 3.1% and 3.9%, respectively. The difference in measurements can be related to the scaling ratio, because in this study the same image processing algorithm was used.
to compute target displacements. Higher pixel number per engineering unit (e.g., millimetre) gives higher measurement accuracy. In this study, one millimetre in S1 and S2 frames at T4 location is approximately 0.06 px (i.e., 0.88 px in 13.3 mm), which is six-time smaller than for GP frames. It is also noticeable that \( \delta_v \) values in Fig. 19 are smaller for the targets with larger \( \delta_v \) such as T4 than for targets with lower \( \delta_v \) such as T1 and T7.

5 Discussion

Semi-automated target detection significantly reduces user input and time. The laboratory beam had painted features (blobs) on its surface. In full-scale bridges, such as the Wilford Suspension Bridge, connections (e.g., hanger to deck connection) can be considered as targets. Machine learning can be employed to automate their detection—similar to what has been achieved in the laboratory study. For user convenience, targets have to have a unique identifier such as a number, which indicates the location of the target on the reference and image planes. For the beam, a sorting algorithm was employed. Targets were first sorted in rows and column, and then a unique number was assigned. When targets with similar features are sought, assigning a ROI for each target helps reducing (i) a likelihood of incorrectly detecting similar targets and (ii) computational time.

The choice and selection of reference points and projection transformation algorithms can be set as an optimization task, in which the set of points providing the highest accuracy are chosen [34]. Selecting a large number of reference points increases a chance of the geometric transformation matrix to become overstrained and provide very accurate results only at no load conditions. The accuracy of the matrix transformation can also be attributed to the accuracy of target locations that are chosen as reference points. The centre of a target could be calculated at a slightly different location than in images taken from different angles, resulting in larger measurement discrepancies. For example, vertical deflections at P6 in Fig. 12 are distinctively different from other camera positions. This can be attributed to setting slightly different coordinates of the reference point for the generation of the planar homography matrix or/and camera side view angle \( \beta \) being significantly larger than for all other camera positions.

5.1 Full-scale applications

There are challenges that need to be addressed for field applications of the approach. Already known issues related to camera drift and stability, and lighting conditions are important. However, it is more important to have a very high measurement resolution. The vertical deflection of the laboratory beam at no damage at its mid-section was 3.3 mm when converted to a convenient form for the assessment of deformation limits, it is the length of the span \( L \) over 330 or \( L/330 \). The vertical deflection serviceability limit states for short to medium span bridges are no larger than \( L/500 \). In normal operational conditions, bridges would seldom have deflections close to their design limits. Therefore high sub-pixel resolution up to 1/500th of a pixel [45] is desirable. Measurement limitations related to resolution can be overcome by reducing the camera field of view and by using multiple cameras such as GoPro connected to synchronization hardware [16]. The measurement accuracy can also be improved using distributed targets of a known pattern [20] and image processing algorithms robust to light-induced image degradation [11]. In such task, the inspecting team need to find the relationship between (i) image resolution, (ii) scale factor (pixel to mm), which is related to the field of view, (iii) and sub-pixel resolution from an image processing algorithm. For example, the horizontal field of view of the GoPro camera is set at 18 m for the Wilford Suspension Bridge monitoring. The camera is set at an angle to the bridge, therefore the closest side (to the camera) has less millimetres per pixel (mm/px) than the far side, which has 3 mm/px (in a vertical direction). Assuming 1/50th of a pixel resolution, which is already high and more realistic than 1/500th of a pixel in the field deployment, gives measurement resolution of 0.06 mm. The required measurement resolution needs to be estimated either before or after the first measurement collection event to define either a suitable field of view or target tracking algorithm.
6 Conclusions

This study introduced a multiple camera position approach for condition assessment of bridges. The premise is that the targets (e.g., surface markers with known dimensions and bridge connections) are located on a single measurement plane, which can be transformed to a 2D reference plane. Movements of targets are tracked when the structure is subjected to known loads (e.g., load truck, train passage). Reference points at a set $x$–$y$ coordinate plane and corresponding points on the structure from a selected image frame are used to generate a geometric transformation matrix, which converts pixels (of targets) to engineering units such as millimetres. Structural response is then computed from target movements at any camera position. The approach is demonstrated on a laboratory beam with artificial targets and a pedestrian suspension bridge with natural targets. Results show that:

- Semi-supervised detection and tracking of targets with known features in a defined region of interest (ROI) for each target provides target locations quickly and accurately. The user has to specify (i) search window of targets in the image, (ii) target features (full circles in laboratory studies) and their corresponding ROI, and (iii) target tracking algorithm.
- 5% discrepancy in computed displacements relative to the mean measurement can be achieved using the geometric transformation at multiple events and multiple camera positions. Such accuracy proved to be sufficient for damage detection and location in the laboratory environment when setting vertical deflections as a damage sensitive parameter.
- The preliminary study on the full-scale bridge demonstrates the capability of the proposed monitoring approach to generate an accurate structural response from multiple camera positions using different cameras and fields of view. The measurement discrepancy in computed displacements is 5.9%. The discrepancy could be reduced by using cameras with zoom lenses (such as GoPro in this study), increasing millimetres per pixels (mm/px) ratio (monitoring part(s) of a bridge) and applying algorithms that offer superpixel resolution.

Measurement discrepancies may increase from camera positions that are significantly different from the initial/reference camera position (such as P6 in Sect. 3.4, see Table 4). A further research is needed to evaluate this statement in a quantitative way. Possibly establishing a training phase for damage identification applications could be included to reduce measurement discrepancy between cameras/camera positions. The developed monitoring approach needs to be further investigated on an event-based measurement collection of full-scale bridges. Bridges that are subjected to known loads such as rail bridges would fit well. Synchronized action cameras with suitable filed of views focusing on small regions of the bridge would give fine measurement accuracy, which should be suitable for the validation of the approach on full-scale bridges. When collecting static response over different seasons, bridge temperature also needs to be measured, even using thermal imaging, to compensate for temperature-induced movements.

Acknowledgements The financial support from Sustainable Futures, Nottingham Trent University, for the equipment and setup used in the laboratory tests and the Wilford Suspension Bridge is gratefully acknowledged. The Wilford Suspension Bridge monitoring was organized by Dr Panos Pismoulis, Mr Sean Ince and Dr Lukasz Bonenberg from Nottingham Geospatial Institute of the University of Nottingham.

Compliance with ethical standards

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Alavi AH, Buttlar WG (2019) An overview of smartphone technology for citizen-centered, real-time and scalable civil infrastructure monitoring. Futur Gener Comput Syst 93:651–672. https://doi.org/10.1016/j.future.2018.10.059
2. Sony S, Laventure S, Sadhu A (2019) A literature review of next-generation smart sensing technology in structural health monitoring. Struct Control Heal Monit 26(3):1–22. https://doi.org/10.1002/stc.2321
3. Kromanis R (2020) Health monitoring of bridges. In: Pacheco-Torgal F, Rasmussen E, Granqvist C-G et al (eds) Start-up creation, 2nd edn. Elsevier Ltd, Amsterdam, pp 369–389
4. Ozer E, Feng MQ (2020) Structural health monitoring. Start-up creation. Elsevier, Amsterdam, pp 345–367
5. Zhao X, Han R, Yu Y et al (2017) Smartphone-based mobile testing technique for quick bridge cable-force measurement. J Bridg Eng. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001011
6. Wang N, Ri K, Liu H, Zhao X (2018) Structural displacement monitoring using smartphone camera and digital image...
7. Zhao X, Ri K, Wang N (2017) Experimental verification for cable force estimation using handheld shooting of smartphones. J Sensors. https://doi.org/10.1155/2017/5625396
8. Xu Y, Brownjohn JMW (2017) Review of machine-vision based methodologies for displacement measurement in civil structures. J Civ Struct Heal Monit. https://doi.org/10.1007/s13349-017-0261-4
9. Dong CZ, Celik O, Catbas NF et al (2019) Structural displacement monitoring using deep learning-based full field optical flow methods. Struct Infrastruct Eng. https://doi.org/10.1080/15732479.2019.1650078
10. Dong CZ, Celik O, Catbas FN et al (2019) A robust vision-based method for displacement measurement under adverse environmental factors using Spatio–Temporal context learning and Taylor approximation. Sensors 19:3197. https://doi.org/10.3390/s19143197
11. Lee J, Lee KC, Cho S, Sim SH (2017) Computer vision-based structural displacement measurement robust to light-induced image degradation for in-service bridges. Sensors (Switzerland). https://doi.org/10.3390/s17102317
12. Zhu J, Lu Z, Zhang C (2020) A marker-free method for structural dynamic displacement measurement based on optical flow. Struct Infrastruct Eng. https://doi.org/10.1080/15732479.2020.1835990
13. Lydon D, Lydon M, Taylor S et al (2019) Development and field testing of a vision-based displacement system using a low cost wireless action camera. Mech Syst Signal Process 121:343–358. https://doi.org/10.1016/j.ymssp.2018.11.015
14. Xu Y, Brownjohn JMW, Huseynov F (2019) Accurate deformation monitoring on bridge structures using a cost-effective sensing system combined with a camera and accelerometers: case study. J Bridg Eng 24:1–14. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001330
15. Ribeiro D, Calçada R, Ferreira J, Martins T (2014) Non-contact measurement of the dynamic displacement of railway bridges using an advanced video-based system. Eng Struct 75:164–180. https://doi.org/10.1016/j.engstruct.2014.04.051
16. Lydon D, Lydon M, Del Rincon JM et al (2018) Development and field testing of a time-synchronized system for multi-point displacement calculation using low-cost wireless vision-based sensors. IEEE Sens J 18:9744–9754. https://doi.org/10.1109/JSEN.2018.285646
17. Dong CZ, Bas S, Catbas FN (2020) A portable monitoring approach using cameras and computer vision for bridge load rating in smart cities. J Civ Struct Heal Monit 10:1001–1021. https://doi.org/10.1007/s13349-020-00431-2
18. Xu Y, Brownjohn J, Kong D (2018) A non-contact vision-based system for multipoint displacement monitoring in a cable-stayed footbridge. Struct Control Heal Monit 25:1–23. https://doi.org/10.1002/stc.2155
19. Feng D, Feng MQ (2016) Vision-based multipoint displacement measurement for structural health monitoring. Struct Control Heal Monit 23:876–890. https://doi.org/10.1002/stc.1819
20. Busca G, Cigada A, Mazzeoli P, Zappa E (2014) Vibration monitoring of multiple bridge points by means of a unique vision-based measuring system. Exp Mech 54:255–271. https://doi.org/10.1007/s11340-013-9784-8
21. Erdenebat D, Waldmann D, Scherbaum F, Teferle N (2018) The deformation area difference (DAD) method for condition assessment of reinforced structures. Eng Struct 155:315–329. https://doi.org/10.1016/j.engstruct.2017.11.034
22. Chu X, Zhou Z, Deng G et al (2019) An overall deformation monitoring method of structure based on tracking deformation contour. Appl Sci. https://doi.org/10.3390/app9214532
23. Kromanis R, Forbes C (2019) A low-cost robotic camera system for accurate collection of structural response. Inventions 4:47. https://doi.org/10.3390/inventions4030047
24. Shao S, Zhou Z, Deng G et al (2020) Experiment of structural geometric morphology monitoring for bridges using holographic visual sensor. Sensors 20:1–25. https://doi.org/10.3390/s20041187
25. Peeters B, De Roeck G (2001) One-year monitoring of the Z24-Bridge: environmental effects versus damage events. Earthq Eng Struct Dyn 30:149–171. https://doi.org/10.1002/1096-9845(200102)30:2<149::AID-EQE1>3.0.CO;2-Z
26. Catbas FN, Susoy M, Frangopol DM (2008) Structural health monitoring and reliability estimation: long span truss bridge application with environmental monitoring data. EngStruct 30:2347–2359
27. Gillins MN, Gillins DT, Parrish C (2016) Cost-effective bridge safety inspections using unmanned aircraft systems (UAS). Geotech Struct Eng Congr 2016:1931–1940. https://doi.org/10.1061/9780784479742.165
28. Feng D, Feng MQ (2017) Experimental validation of cost-effective vision-based structural health monitoring. Mech Syst Signal Process 88:199–211. https://doi.org/10.1016/j.ymssp.2016.11.021
29. Klaiber FW, Wipf TJ, Kempers BJ (2003) Repair of damaged prestressed concrete bridges using CFRP. In: Proceedings of the 2003 Mid-Continental Transportation Research Symposium. Ames, Iowa
30. Lee JH, Cho S, Sim SH (2015) Monocular vision-based displacement measurement system robust to angle and distance using homography. In: Int Conf Adp Exp Struct Eng Aug 2015
31. MathWorks (2020) Image Processing Toolbox: User’s Guide (R2020b). Retrieved November 14, 2020 from https://nl.mathworks.com/help/pdf_doc/images/images Ug_pdf
32. Harris CG, Stephens M (1988) A combined corner and edge detector. Alvey Vis Conf 15:10–5244. https://doi.org/10.5244/C.2.23
33. Bay H, Tuytelaars T, Van Gool V (2008) SURF: speeded-up robust features. Eur Conf Comput Vis. https://doi.org/10.1007/11744_023_32
34. Kromanis R, Liang H (2018) Condition assessment of structures using smartphones: a position independent multi-epoch imaging approach. In: 9th European Workshop on Structural Health Monitoring Series (EWSHM). Manchester, UK
35. Kromanis R, Xu Y, Lydon D et al (2019) Measuring structural deformations in the laboratory environment using smartphones. Front Built Environ. https://doi.org/10.3389/fbuil.2019.00044
36. Mei Q, Gül M (2018) A crowdsourcing-based methodology using smartphones for bridge health monitoring. Struct Heal Monit. https://doi.org/10.1177/1475921718815457
37. Feng D, Feng MQ (2016) Output-only damage detection using vehicle-induced displacement response and mode shape curvature index. Struct Control Heal Monit 23:1088–1107. https://doi.org/10.1002/stc
38. Bakhtiar-Nejad F, Rahabi A, Esfandiarib A (2005) A structural damage detection method using static noisy data. Eng Struct 27:1784–1793
39. Qiao P, Lestari W, Shah MG, Wang J (2007) Dynamics-based damage detection of composite laminated beams using contact and noncontact measurement systems. J Compos Mater 41:1217–1252. https://doi.org/10.1177/0021998306067306
40. Dackermann U, Li J, Samali B (2010) Dynamic-based damage identification using neural network ensembles and damage index method. Adv Struct Eng 13:1001–1016. https://doi.org/10.1269/journal.4332.13.6.1001
41. Dong CZ, Celik O, Catbas FN et al (2020) Structural displacement monitoring using deep learning-based full field optical flow methods. Struct Infrastruct Eng 16:51–71. https://doi.org/10.1080/15732479.2019.1650078
42. Yuen HK, Princen J, Dlingworth J, Kittler J (1990) A comparative study of hough transform methods for circle finding. Image Vis Comput 8:71–77. https://doi.org/10.5244/c.3.29

43. Kromanis R, Al-Habaibeh A (2017) Low cost vision-based systems using smartphones for measuring deformation in structures for condition monitoring and asset management. In: 8th International Conference on Structural Health Monitoring of Intelligent Infrastructure. Brisbane

44. British Standards Institution (2003) BS EN 338:2016: Structural timber—Strength classes. London, British Standards Institution

45. Imetrum (2020) Digital image correlation. https://www.imetrum.com/products/digital-image-correlation/. Accessed 2 Feb 2020

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.