Real-Time Vision-Based Dynamic Monitoring of Vertical and Horizontal Transverse Displacements in Bridge Decks

Fabio Micozzi[1]; Michele Morici[2]; Alessandro Zona, M.ASCE[3]; and Andrea Dall'Asta[4]

Abstract: The objective of this study is to present a vision-based system that can extract in real-time displacements at multiple cross sections of bridge decks, which is accurate, easily replicable, and cost-effective. The key feature to accomplish such a goal is the reduction of the computational time required for displacement extraction, achieved through the use of multiple image subdomains whose position is continuously updated in the acquired video footage, following the movements of the monitored targets. This optimization at the software level is paired with choices in the hardware configuration that provide major benefits, namely: (1) camera below the bridge deck with its optical axis parallel to the bridge axis to obtain vertical and horizontal transverse displacements of any cross section, even with a single video camera; and (2) use of coded targets for automated camera calibration, automated definition of the initial positions of the subdomains, and automated compensation of the vibrations of the video camera employing a reference point in the abutment or intermediate pier that naturally falls within the camera field of view. An application to a post-tensioned concrete bridge during normal traffic conditions is used to analyze the quality of the obtained results through comparisons with contact sensors (displacement transducer and accelerometer). The results show that the proposed computer vision system meets the declared research objectives, given its precision, even in the case of displacements of a few millimeters, as in the considered case study, broad applicability for the most common bridge decks, and the optimized use of affordable hardware. **DOI: 10.1061/JBENF2.BEENG-7597**.

Author keywords: Bridge monitoring; Computer vision; Digital image correlation; Experimental analysis; Fiducial targets; Structural health monitoring; Target matching; Vibration analysis; Vision-based monitoring.

Introduction

Structural health monitoring (SHM) is an essential tool for ensuring the safety of bridges during their lifespan and planning condition-based maintenance, as widely acknowledged in the engineering community (Wenzel 2009; Farrar and Worden 2013; Webb et al. 2015; Chen 2018). Evolutions and advancements in SHM are inevitably correlated to the available sensor technologies that determine which structural response parameter can be measured and the accuracy of its measure. Among the structural response parameters, displacements are generally considered the most important ones because they provide valuable information in a wide range of conditions, from slow quasi-static movements (settlements in the foundations, time-dependent deformations in concrete and tendons, thermal effects, dislocations of bearing devices, etc.) to the dynamic response due to vehicular traffic, wind, or earthquakes. For example, a recent study relying on the Hilbert–Hung transform (Scozzese and Dall'Asta 2024; Scozzese et al. 2025) showed that displacement measurements can be conveniently used to extract the instantaneous properties of the dynamic response, with application to the characterization of the nonlinear behavior of post-tensioned RC bridges under vehicular traffic.

The identification of the structural response of a bridge deck requires that both vertical and horizontal components of the displacements are simultaneously measured, in multiple cross sections, given the contributions to the overall movements from vertical bending, torsion, transverse bending, and distortion. Undeniably, measuring vertical and horizontal transverse displacements of bridge decks is a challenging task. Displacement transducers (DTs) cannot be used due to the lack of stationary points close to the deck, except for those few decks with limited height above the ground or for special situations where stiff auxiliary structures are built to support the transducers. Other contact sensors can be exploited for an indirect determination of displacements. For example, inclinometers and strain gauges provide local measures that can be used to derive displacements (Hou et al. 2005; Zhang et al. 2017; Ozdagli et al. 2017; Sun et al. 2020), but they require a structural model that inevitably influences the results. Accelerometers, probably the most common sensors in SHM, can be used to obtain displacements (limited to dynamic loading conditions only) from double integration of accelerations (Park et al. 2005; Gindy et al. 2008; Moreu et al. 2016; Sekiya et al. 2016; Bunce et al. 2023), but special care is necessary to avoid inaccuracies introduced in the numerical integration process and baseline corrections.

[1]Postdoctoral Engineer, Scientific and Technical Laboratory Support Area, Univ. of Camerino, Via Le Mosse 20, Camerino 62032, Italy. ORCID: https://orcid.org/0000-0001-8651-670X. Email: fabio.micozzi@unicam.it

[2]Associate Professor, School of Architecture and Design, Univ. of Camerino, Viale della Rimembranza 3, Ascoli Piceno 63100, Italy. ORCID: https://orcid.org/0000-0001-9032-388X. Email: michele.morici@unicam.it

[3]Professor, School of Architecture and Design, Univ. of Camerino, Viale della Rimembranza 3, Ascoli Piceno 63100, Italy (corresponding author). ORCID: https://orcid.org/0000-0002-3100-1802. Email: alessandro.zona@unicam.it

[4]Professor, School of Science and Technology, Univ. of Camerino, Via Gentile da Varano III 26, Camerino 62032, Italy. https://orcid.org/0000-0001-7482-9434. Email: andrea.dallasta@unicam.it

Note. This manuscript was submitted on January 24, 2025; approved on July 14, 2025; published online on September 26, 2025. Discussion period open until February 26, 2026; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Bridge Engineering***, © ASCE, ISSN 1084-0702.

Interesting alternatives in bridge monitoring are offered by contactless technologies (Narasimhan and Wang 2020), and among them, vision-based monitoring appears as the simplest and cost-effective solution for direct measurement of displacements in quasi-static and dynamic conditions (Catbas et al. 2012; Alipour et al. 2019; Dhanasekar et al. 2019; Xu et al. 2019; Ghyabi et al. 2023; Zhou et al. 2023; Debees et al. 2024). Comprehensive state-of-the-art reviews on the application of computer vision to monitoring displacements in bridges and other structures can be found in Feng and Feng (2018), Dong and Catbas (2021), Zona (2021), and Luo et al. (2023). The underlying concept is simple: the displacements are obtained from the analysis of subsequent images in a video footage. Advanced image-processing algorithms can achieve subpixel resolution in the computation of displacements, thus, prospecting appealing potentialities for monitoring civil structures (Feng and Feng 2021). However, differently from the use of contact sensors, relying on consolidated procedures for installation and acquisition, vision-based monitoring is a younger technology with no established protocols. Accordingly, effective indications on specific aspects, from software for displacement extraction to hardware configuration, are valuable information toward the future definition of guidelines for computer vision applications in bridge monitoring.

The objective of this study is to present a vision-based system that can extract displacements at multiple cross sections of bridge decks. The system is designed to be accurate, easily replicable in a wide variety of situations, and cost-effective. The key implemented feature is the use of multiple image subdomains for each of the monitored targets; the positions of the subdomains are continuously updated in the acquired video footage to reduce the computational time of the algorithm for displacement extraction. This optimization at the software level has two major implications: (1) possibility to monitor the structural dynamics response in real time at acquisition frequencies much higher than the bridge fundamental frequency, with no need to store memory-consuming videos for later postprocessing, as is instead the case of most of the applications found in the technical literature; and (2) possibility to extract more displacement components from the same video camera, which, in turn, simplifies hardware configuration and installation, reduces costs, allows real-time corrections of the disturbances induced in the camera, thanks to one or more stationary points within the field of view (FOV) of the video camera.

The proposed optimization at the software level is paired with choices in the configuration of the hardware (video camera and targets), listed as follows: (1) a video camera below the bridge deck with optical axis parallel to the bridge axis to obtain the vertical and horizontal transverse displacements at any cross section, even with a single camera; and (2) coded targets for automated camera calibration, automated definition of the initial positions of the subdomains, and automated correction of disturbances using a point, on the opposite pier or abutment, that is expected to be always within the camera FOV. Such a configuration can be replicated to any rectilinear bridge with accessible abutments and pier caps. Attention must be given to the fact that the proposed configuration requires the use of algorithms for displacement extraction that must be insensitive to targets being partially out of focus, an inevitable condition, because targets at different cross sections of the bridge deck have different distances from the video camera.

The proposed vision-based system, implemented through the combination of MATLAB codes (MathWorks 2024) and affordable hardware, is applied to an existing in-service post-tensioned concrete highway bridge. The peculiar characteristics of the adopted case study are the possibility to directly compare vision-based and contact measurements of displacements, thanks to the limited distance of the bridge midspan from the ground. This constitutes a special condition that allows the assessment of the precision that can be achieved with the adopted hardware and software. It should be remarked that the considered case study is a test more severe than previous applications in the literature, such as steel bridges and footbridges characterized by larger displacements and lower frequency contents.

Workflow for Vision-Based Dynamic Monitoring of Displacements

Moving Subdomain

The common workflow for displacement tracking using computer vision starts with the preliminary identification of a region of interest (ROI) within the FOV where an algorithm searches for the position of the monitored target. Hence, the ROI must contain the initial position of the target and all its subsequent positions. Accordingly, an assumption must be introduced on the expected amplitude of the displacements. If the displacements are outside the assumed range, then the procedure fails to provide the correct measurements; if the displacements are much smaller than assumed, then the procedure is computationally inefficient because the tracking algorithm is forced to work on the entire ROI, despite the targets remaining in a small portion of the ROI.

A different approach is proposed in this study to avoid that the initial guess on the amplitude of displacements might have a negative impact on the accuracy and efficiency of the adopted computer vision monitoring of displacements. Given that in structural monitoring, small movements of the target are generally expected between two subsequent images with a temporal distance of less than a tenth of a second, in this study, an optimization is proposed: the search for the new position of the target can be conveniently made within a smaller area (sub-ROI) included in the ROI. While the ROI is fixed and its position is defined at the beginning of the acquisition, the sub-ROI is defined at the software level and its position is continuously updated.

The concept is that the monitored target has an associated sub-ROI that is centered on the target at a given frame and has sufficient margin to contain the position of that same target in the subsequent frame. The position of the sub-ROI is recentered on the target frame after frame, following its movement. In this way, the computation cost of the proposed workflow is strongly reduced as the algorithm switches from a global search within the ROI to a local search within the smaller sub-ROI, being the analysis time proportional to the area where the search is performed. Accordingly, the initial selection of a larger ROI does not have negative implications for computational efficiency.

Combined Template Matching and Windowing

Different computer vision algorithms for displacement tracking are known for their accuracy and computational efficiency (Feng and Feng 2021), which can be used in conjunction with the proposed sub-ROI approach. The algorithm used in this study for the vision-based measure of displacements is template matching through an area-based method, adopting an upsampled cross-correlation by means of a fast Fourier transform (FFT) in the image spatial domain, as developed by Guizar-Sicairos et al. (2008). It achieves subpixel resolution with small computational time and has reduced memory requirements.

In this study, the performance of the algorithm by Guizar-Sicairos et al. (2008) is further improved according to the

suggestions made by Stone et al. (2003) to reduce the noise introduced by FFT matching for both high and low frequencies. First, a Tukey windowing function is applied to avoid the phenomenon known as spectral leakage, where the truncated signal shows up in the FFT as high-frequency components not in the original signal. The residual aliasing effects in the low frequencies of the FFT are afterward corrected by applying a high-pass filter in the FFT of the image. The final effect of windowing can be visualized as a shading of the outer regions of the image, with the pixels at the border being set to zero (black).

The negative effects of spectral leakage are particularly significant when processing small images, as is the case with the ROI subdomains implemented in this study. This condition makes the implemented windowing procedure an essential tool for displacement extraction in the presented computer vision workflow, with major improvements in terms of noise reduction, robustness, and repeatability of the displacement measurements. Significant inaccuracies or even an incapacity to extract displacements were observed when tracking operates on small images without windowing. While such benefits were acknowledged in the computer vision technical literature, no previous applications in vision-based structural monitoring could be found.

Multiple Subdomains for Handling Multiple Targets

In the present study, multiple targets acquired by the same video camera are considered. As already discussed, in this way it is possible to (1) extract absolute displacement time histories of multiple points, (2) obtain relative displacements between the considered points, and (3) compensate the effects of the disturbances in the video camera if there is at least a stationary point among the monitored ones. The first step is the selection of an ROI within the FOV, containing the targets whose displacements will be monitored (Fig. 1). The targets must not be superimposed in their initial position as well as during their movements and must always be contained in the ROI. Afterward, for each target, a subset of the ROI containing only a single target is defined (sub-ROIs as depicted in Fig. 1). In this way, a significant reduction of the search area can be achieved when considering multiple targets to benefit computational efficiency.

Hardware Configuration Oriented to Bridge Deck Monitoring

Applications of computer vision to bridge monitoring generally use video cameras installed in positions external to the footprint of the bridge plan, typically pointing at targets or structural features on the sides of the bridge deck (Feng and Feng 2018; Dong and Catbas 2021; Zona 2021; Luo et al. 2023). Hence, no distinction between torsional and flexural components can be made using a single camera. In addition, relying on lateral views of the bridge does not constitute an easily repeatable protocol given that the bridge side areas could be inaccessible in various situations, e.g., presence of a river or a lake, impervious mountain areas, interferences with other roads, or interdicted access to private properties. Very few exceptions can be found; for example, Aliansyah et al. (2021) installed their video camera at the road level with optical axis parallel to the outer edge of the guardrail, while Yin et al. (2023) installed a network of cameras inside a steel box deck, in both cases to monitor the trend of the vertical deflection along the entire bridge.

A different configuration is proposed in this study (Fig. 2): the video camera is placed under the bridge deck, on an abutment, or on top of a pier, with its optical axis parallel to the bridge axis. This appears as the most convenient position, given that (1) it is always present in a bridge and expected to be accessible, for example, for inspecting bearings and joints; (2) it allows to obtain both vertical deflections and horizontal transverse displacements at any cross sections from a single camera; (3) the camera is in a position protected from direct sunlight, atmospheric events, and not interfering with traffic, making such vantage point a convenient choice also for permanent monitoring; and (4) stationary points at the end of the span, e.g., opposite abutment or pier, fall within the camera FOV.

The position of the camera under the bridge deck is complemented by coded targets for automated camera calibration and target identification, installed at the intrados of the deck with their plane orthogonal to the bridge axis. The use of targets fosters measurement replicability because targetless monitoring is not always a possible option in the case of an optical axis parallel to the deck axis, especially for some structural typologies with uniform surfaces. In addition, coded targets provide the possibilities for automated camera calibration, automated definition of the initial position of the sub-ROIs, and automated compensation of the movements of the video cameras, as detailed in the following paragraphs.

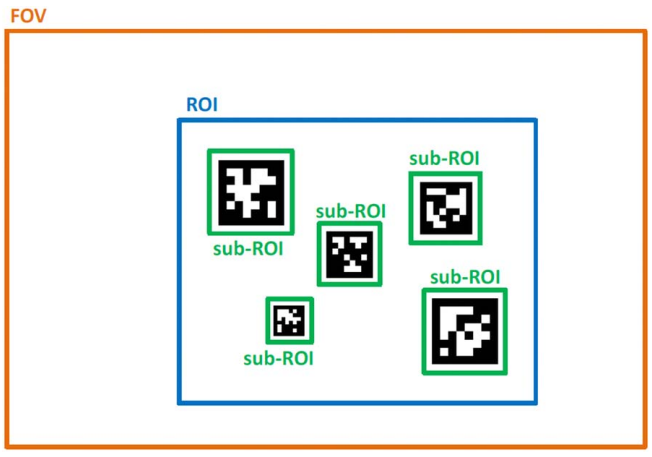

Fig. 1. Exemplification of the implemented multiple sub-ROIs.

Depth of Field and Out-of-Focus Influence

It should be remarked that in the proposed hardware configuration, the targets in different cross sections have different distances from the video camera. Accordingly, only the targets of a selected cross section, e.g., midspan, can be precisely focused, while all other targets at different cross sections, e.g., other than midspan, will be out of focus.

The amount of blurring of the out-of-focus targets depends on the distance from the focused cross section and is influenced by the hardware depth of field, i.e., the distance between the nearest and the farthest objects that are in acceptably sharp focus. In photography, the depth of field is commonly controlled through lens aperture; e.g., reducing aperture diameter increases the depth of field, with the downside that the quantity of light reaching the camera sensor is reduced, requiring longer exposure time. While this simple rule can be conveniently exploited when setting the proposed hardware configuration, it is very important to consider that blurring has a small impact on the measured displacements when area-based methods are used for displacement extraction, as is the case with the proposed workflow. This condition holds inside a large acceptable range, as long as blurring leaves the target distinguishable, i.e., the sequences of transitions between darker

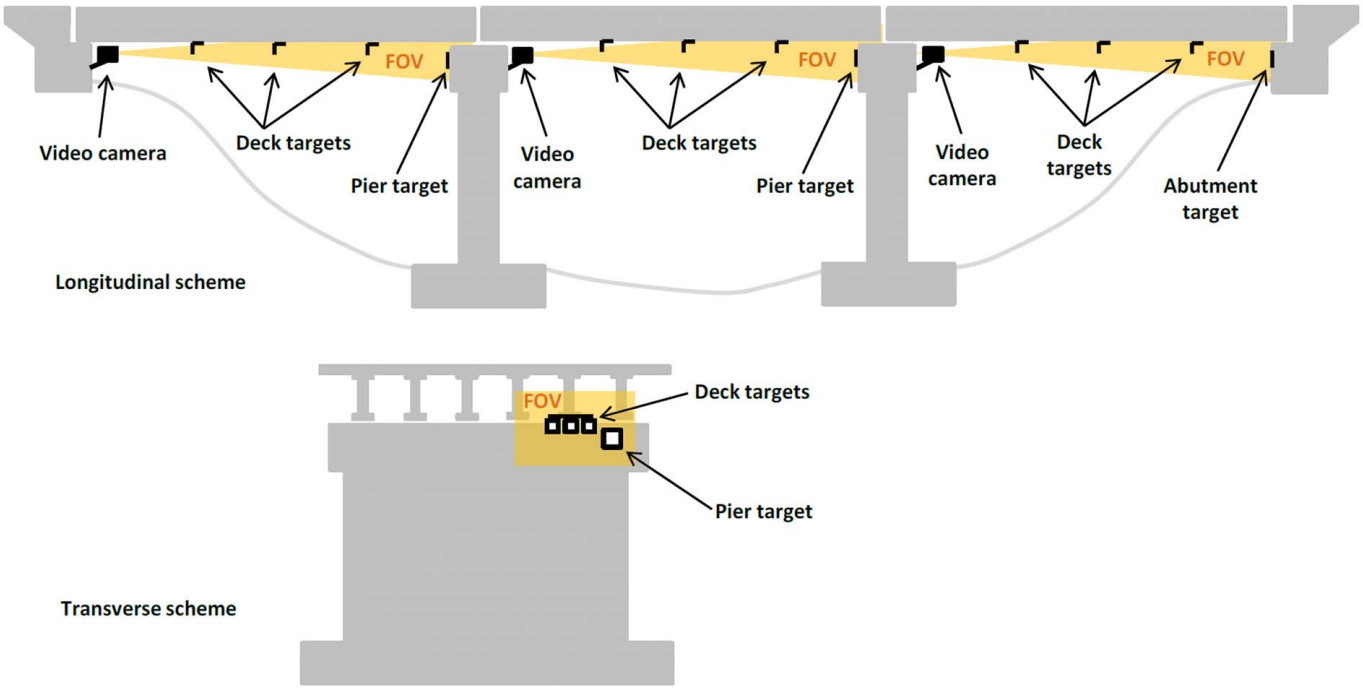

Fig. 2. Exemplification of the proposed hardware configuration.

and lighter portions of the target chessboard geometry are preserved even if their contour is no longer sharp.

Conversely, other algorithms for displacement extraction, such as feature-based methods, are negatively influenced by blurring and might introduce significant errors and noise when targets are out of focus. In this case, a solution might be the use of more complex optical configurations where points at different distances can be equally focused (Xing et al. 2022).

Simplified Camera Calibration

The procedure that calculates the physical displacement of the monitored target from pixel displacements acquired by the video camera sensor is called camera calibration. If the camera optical axis is perpendicular to the displacement plane, the calibration reduces to a simple scale factor. The determination of the scale factor ratio R_{SF} is based on geometric relations that provide

$$R_{SF} = \frac{d_0}{d_0^i / d_{pixel}} = \frac{D}{f} d_{pixel} \tag{1}$$

where d_0 = physical dimension of the target; d_0^i = corresponding physical dimension on the sensor plane of d_0; d_{pixel} = pixel size (physical characteristic of the sensor of the video camera); D = distance between the camera focal point and the target; and f = focal length of the lens (Fig. 3). It is observed that Fig. 3 was depicted without respecting the actual proportions because the focal length and the projection of the target on the sensor are much smaller than the target distance and dimension, respectively. Although the condition of an optical axis perpendicular to the structural surface is very difficult to satisfy in many real conditions, Feng and Feng (2021) reported negligible errors in displacement estimation in practical applications, e.g., about 1% error for a tilt angle of

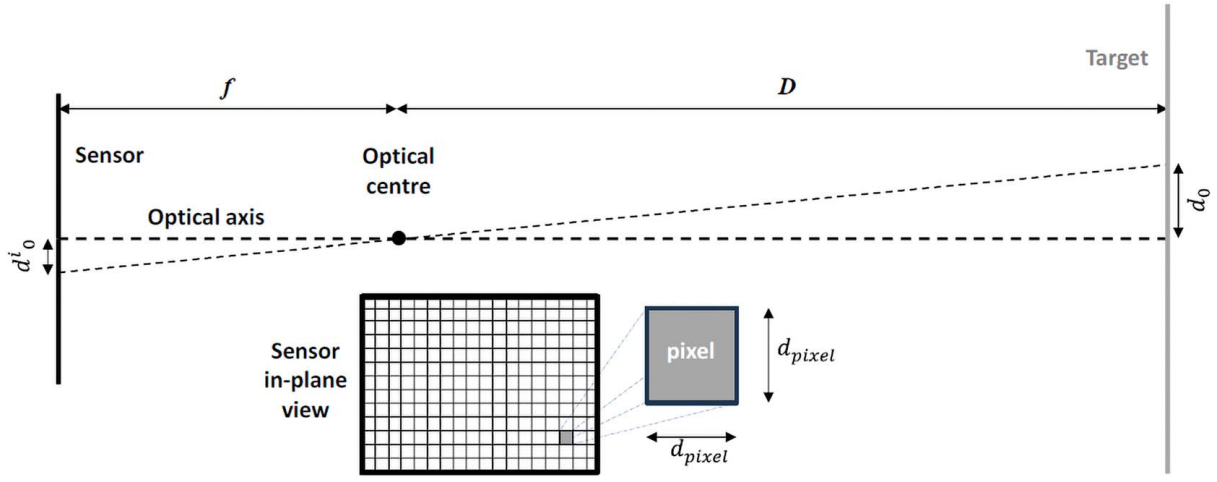

Fig. 3. Sensor–target optical scheme.

30° when using a 50 mm focal length, which decreases when the focal length increases.

Procedure for Camera Calibration and Subdomain Definition

Commonly, R_{SF} is computed by counting the number of pixels for a known dimension of the target. Such a procedure can be automated through the use of fiducial targets called AprilTags (Olson 2011, Wang and Olson 2016, Krogius et al. 2019), allowing their automatic identification with minimal computational overhead. The tags consist of a square grid of black and white cells, with a distinctive pattern that allows for robust detection and identification even under challenging conditions and blurred (partially out-of-focus) images. Among the different AprilTag families, the 36h11 family, based on a 6×6 grid of data cells surrounded by a black border (examples are reported in Fig. 1), is adopted in this work. These tags can be used with the AprilTag detection software (available in MATLAB), which computes the precise three-dimensional position, orientation, and identity of the tags relative to the camera.

The scale factor based on the pixel-to-real dimension of the tags is automated using the average length value between the four corners of the target as detected by the AprilTag software with a subpixel accuracy, knowing the real dimension of each identified target. In this way, the error introduced by a manual pick-up of the points on the frame is eliminated, increasing the speed of the set-up and the repeatability of the measures.

It should be remarked that the AprilTag search algorithm is not fast enough to work in real-time for the frequency required in structural dynamic monitoring of displacements, typically equal to or larger than 50 frames per second (FPS) for medium- and short-span bridges. Consequently, the AprilTag code is used in this study to detect the multiple targets visible in the ROI only during the system set-up to automatically compute the scale factors for each target and automatically define the relevant position of the sub-ROIs.

Procedure for Compensating Camera Movements

Vision-based displacement measuring is always affected by a possible source of noise related to the motion of the camera, resulting in an apparent motion of the targets that could reduce the measure precision. The proposed installation under the bridge exposes the video camera to the disturbance of the nearby traffic. Consequently, a strategy to compensate for the camera motion should always be considered. There are different methods in the technical literature for camera motion correction, as discussed in Feng and Feng (2021). The simplest one is based on the acquisition of a reference stationary target in the same FOV as the considered targets. The apparent motion of the stationary reference target can be assumed as determined only by the vibration of the camera, allowing its removal (Yoneyama and Ueda 2012, Feng and Feng 2017).

Some clarifications on the camera movement correction are reported hereafter. A Cartesian reference system with origin at the center of the camera sensor defines the plane of the sensor, whereas the third axis defines the optical axis. The six components of the possible movement of the sensor are, accordingly, two in-plane translations, one out-of-plane translation, and the three rotations (pitch, yaw, roll). The in-plane translation vector equal to $\boldsymbol{v_c}$ results in an apparent displacement of the target of the same magnitude and opposite direction and can be computed from simple geometric relations derived from Fig. 4(a)

$$(\boldsymbol{d}_1^i - \boldsymbol{d}_0^i) R_{SF} = \boldsymbol{d}_1 - \boldsymbol{d}_0 = \boldsymbol{v_c} \tag{2}$$

Consequently, the effect of the camera translation in the sensor plane is significant if $\boldsymbol{v_c}$ is of the same order of magnitude as the displacement measurement precision.

The effect of pitch or yaw rotation is more complex [Fig. 4(b)]; however, under the hypothesis of small rotation, the distance of the center of the sensor from the target plane can be considered equal for all the points of the sensors, and the apparent displacement of the target can be computed as follows:

$$(\boldsymbol{d}_1^i - \boldsymbol{d}_0^i) R_{SF} = \boldsymbol{d}_1 - \boldsymbol{d}_0 = (f + D) \tan(\boldsymbol{\alpha_c}) \tag{3}$$

where $f + D$ = total distance of the sensor from the target; and $\boldsymbol{\alpha_c}$ = vector containing pitch and yaw rotations.

The out-of-plane translation (movement of the camera along the optical axis) changes the distance D between the sensor and the target, producing an apparent movement through a modification of the scale factor R_{SF}. However, such translational movement is much smaller (various orders of magnitude) compared to the sensor–target distance D and its effect is accordingly negligible. The effect of roll rotation (around the optical axis) is of an order of magnitude lower than the effect of the in-plane translations and out-of-plane rotations; hence, it is neglected in the adopted corrections of the video camera movements. Accordingly, the camera motion can be described as a combination of a translation $\boldsymbol{v_c}$ and rotation $\boldsymbol{\alpha_c}$. The apparent motion of a reference target is, hence, detected by the sensor for both translation $\boldsymbol{v_c}$ and rotation $\boldsymbol{\alpha_c}$ through the previous highlighted difference $\boldsymbol{d}_1^i - \boldsymbol{d}_0^i$. Consequently, although it is not possible to identify the weight of the two contributions, the error correction for the camera motion can be carried out simply by removing the pixel displacement of the reference target from the acquired pixel displacement of the targets.

Application to a Medium-Span Post-Tensioned Concrete Highway Bridge

Case Study

The considered case study is a three-span concrete bridge located in central Italy, part of the national highway road network. Each span has a single deck (dual carriageway with two lanes per direction) made by a concrete slab on six simply supported post-tensioned concrete beams (span length 32 m) connected by end and intermediate transverse beams at about 1/3 and 2/3 of the span. The West span of this bridge was selected as a case study, given its reduced distance from the ground at the midspan, making possible a simple installation of displacement transducers for comparative purposes. A first experimental campaign on this bridge using computer vision was conducted on July 20, 2023; the outcomes were reported by Micozzi et al. (2023, 2024). A second experimental campaign, made to provide information for the discussion of the improvement points introduced in this study, was performed on May 10, 2024. In both experimental campaigns, the bridge was normally operational with regular weekday vehicular traffic, mostly cars and less frequent passages of heavy trucks. It is remarked that the novelties, introduced in this second experimental campaign with respect to first one, are (1) improved image-processing algorithm with the sub-ROI procedure for optimized displacement tracking; (2) multiple targets, including one assumed as stationary for compensation of the vibrations in the video camera, with real-time simultaneous extraction of both vertical and horizontal transverse displacements; and (3) AprilTags for

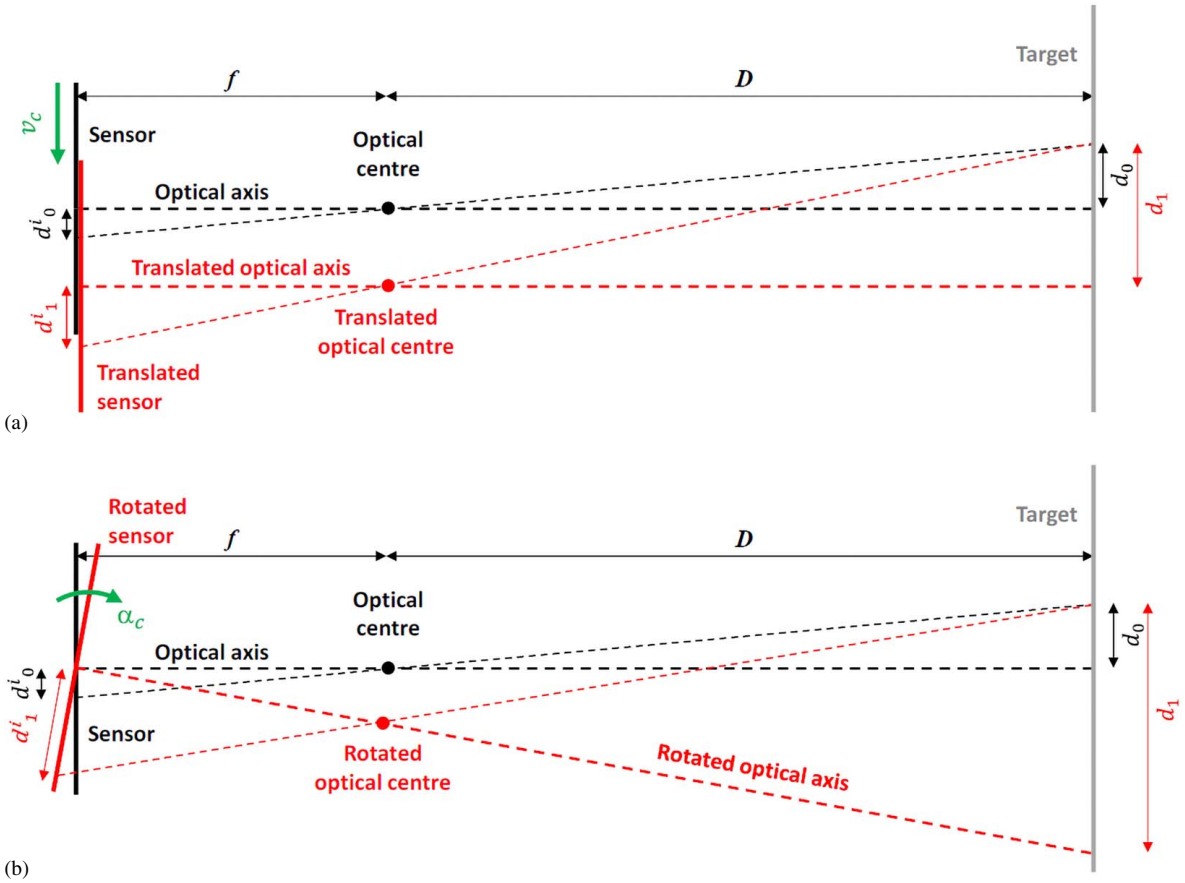

Fig. 4. Effects of the main movements of the sensor: (a) effect of sensor in-plane translation; and (b) effect of sensor out-of-plane rotation.

automated target identification, camera calibration, and definition of the initial position of the sub-ROIs.

Computer Vision Equipment and Settings

An industrial computer vision video camera with a resolution of 1,920 × 1,200 pixels (Teledyne FLIR Blackfly S BFS-U3-23S3M-C, Wilsonville, OR) connected to a laptop (Dell Precision 7560, Round Rock, TX) through USB3.0 was placed under the bridge deck, near the West abutment, under the second beam on the south side of the deck (Fig. 5) on a professional-grade aluminum video tripod. The video camera was equipped with a lens (Kowa LM100JC1MS, Nagoya, Japan) having a fixed focal length $f = 100$ mm, and the aperture was set at $f/8$ as a balance between extension of the depth of field and amount of light reaching the camera sensor for the selected shutter speed. Detailed specifications of the adopted hardware can be found in Micozzi et al. (2023). Electrical power supply was provided by a portable solar generator.

The video camera is placed about 60 cm below the beam extrados with its optical axis essentially parallel to the bridge axis, having a tilt angle of about 2°, much lower than previously mentioned tilt values reported to cause negligible errors in displacement

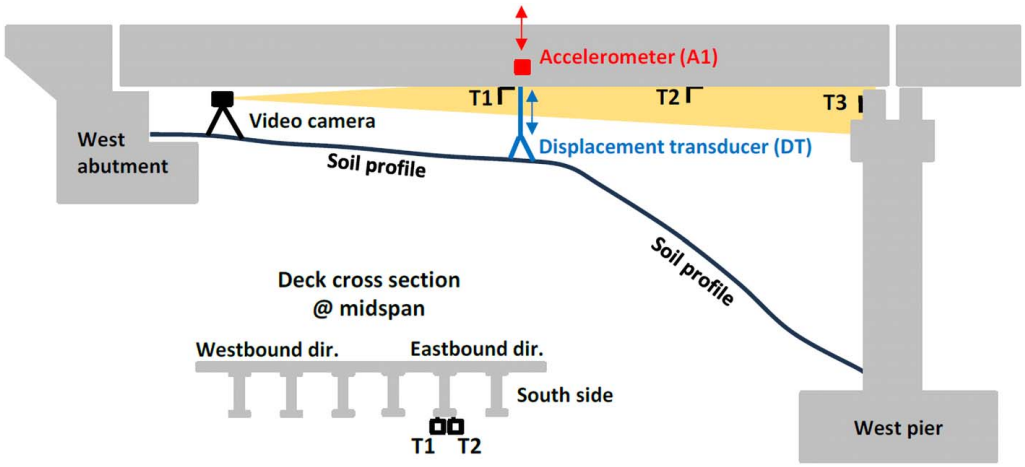

Fig. 5. Geometry of the considered case study with the schematic position of the monitoring equipment (video camera, targets, contact sensors).

Fig. 6. View of the considered case study: (a) pier as seen under the monitored span with indication of the position of the assumed stationary target T3; (b) detail of the pier cap with Target T3; (c) view of the three targets (T1 at the midspan, T2 at the 3/4 span, and T3 at the pier), DT and accelerometer (A1) at the midspan; and (d) video camera and acquisition system.

estimation (Feng and Feng 2021). Perspective distortion in the adopted configuration is extremely low, given the small tilt angle of the optical axis and the fact that displacements are very small compared to the distance between the camera and targets. Specifically, the camera paired with the adopted lens has a FOV with an aperture of about 2.4° vertically and 3.2° horizontally, the entire ROI has nearly 1° of aperture, and the displacements recorded are <5 mm (0.05° in terms of angle). In addition, the lens has a very small barrel distortion (0.05%). Accordingly, optimal optical specifications are assured in the adopted hardware installation.

Targets were aluminum composite panels with a white surface and black matte film shaped using high-precision laser cut, according to the AprilTag design. Two targets, T1 and T2, were installed at the extrados of the second beam using a steel L-shaped element fastened to the concrete, T1 at the midspan and T2 at the 3/4 span. A third target was installed at the pier cap, directly fastened to the concrete. Fig. 6 shows some views of the monitored span and instrumentations, while Fig. 7 shows the targets as viewed by the video camera. Targets T1 and T2 have the same dimensions (the side of the black square is 80 mm), and their AprilTag IDs are 4 and 5, respectively. Target T3 is double the size of T1 and T2; i.e., 160 mm is the side of its black square, and its AprilTag ID is 11. The targets, having different distances from the video camera, cannot be simultaneously focused. However, as already discussed, the adopted area-based method for displacement tracking is weakly influenced by blurred targets as long as they are recognizable; this condition was verified by the experimental results presented afterward.

The acquisition frequency was set at 60 FPS, and the exposure time was set at 0.005 s. The value of the upsampling factor was set equal to 50, thus theoretically providing a resolution equal to 1/50 of a pixel. However, this is the maximum expected resolution, not the accuracy of the measure that is afterward estimated from the results of the experimental field application, inevitably influenced by mechanical, optical, and electrical noises in the adopted system and its interactions with the external environment.

The selected ROI was a square window of side 700 pixels; the sub-ROIs were square windows of side 236, 162, and 204 pixels for Targets T1, T2, and T3, respectively. The margin between the target and the sub-ROI was set to 1/8 the size of the target (10 mm for Targets T1 and T2, 20 mm for Target T3), identified as the optimal value in the application presented. The resulting scale factors were 0.424, 0.617, and 0.980 mm/pixel, providing theoretical resolutions (computed as 1/50 of a pixel given the adopted upsampling factor) equal to 0.0085, 0.0123, and 0.0196 mm for Targets T1, T2, and T3, respectively. It is observed that the image resolution of the targets, i.e., their dimension in pixels, is larger than the resolution (square with side of 100 pixels) indicated in Micozzi et al. (2023) as the optimal value for target tracking based on experimental field tests. This was achieved using a larger size for the target that is more distant (target T3).

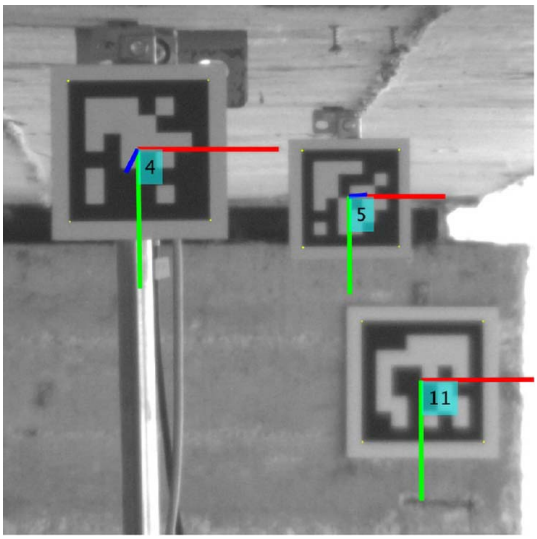

Fig. 7. Targets as viewed by the video camera within the selected ROI and recognized by the automated AprilTag identification and calibration (T1 = #4, T2 = #5, T3 = #11).

The use of sub-ROIs determined a reduction in the elaboration time of about 20 times as compared to target search over the entire ROI, i.e., the computational time required to elaborate each video frame was reduced from 0.25 s (without subdomains) to 0.0125 s (with subdomains), resulting in an increment from 4 to 80 FPS of the maximum acquisition frequency compatible with real-time processing. More in detail, the processing time of 0.0125 s for each acquired image frame is subdivided as follows: image transfer from the camera to the laptop, 0.001 s; initial cross-correlation, 0.002 s; upsampling cross-correlation, 0.009 s; and video plot of displacement computed and data management, 0.0005 s.

Vertical and horizontal displacements of the three targets were extracted in real time and stored as text files. In addition, the raw videos of the ROI were saved as backup files for possible subsequent evaluations using processing settings different with respect to those adopted in the real-time elaboration.

Installed Contact Sensors

A high-precision linear variable displacement transducer (Gestecno TSL-160, Castelraimondo, Italy) measuring the vertical displacements between the extrados of the second beam and the ground was installed at the midspan, as allowed by their reduced distance (about 1.8 m). A high-sensitivity piezoelectric accelerometer (PCB 393B31, New York) measuring the vertical acceleration was also installed at the midspan in the extrados of the second beam. The signals of the displacement transducer and accelerometer were acquired using high-performance data logging hardware (Dewesoft Krypton STG6 and Krypton ACC8, Trbovlje, Slovenia) and its controlling software, with the acquisition sampling frequency set to 1,200 samples/s. The same solar generator that provided electrical power to the computer vison set-up also alimented the contact-based monitoring system. In addition, a sensor for air temperature and relative humidity (RH) with an integrated data logger (Elitech RC-51H, San Jose, CA) was placed at the midspan, on the lower flange of the monitored beam.

Testing Sequences and Environmental Conditions

During the experimental campaign, made on May 10, 2024, between 1:00 and 4:30 p.m., repeated 10-min recordings were made. Selected results are afterward reported and discussed for the last measurement window that started at 4:20 p.m., being the one with the largest vehicular traffic. The weather was sunny. Temperature and humidity were almost constant during the entire monitoring campaign, ranging between 20°C and 21°C and 54% and 59%, respectively.

Vertical Displacements without Camera Movement Compensation

Fig. 8 shows the vertical displacement in pixels of the three targets, i.e., the displacement projected on the camera sensor, before the application of the scale factor R_{SF} to convert the pixel measure to the actual target displacement. The in-plane translation and pitch rotation of the video camera sensor induce the same pixel (apparent) displacement for the three targets. The measurements show the passage of each vehicle and the fluctuations of T3 (that can be considered a stationary point) caused by the vibrations induced in the video camera by the traffic. Some peaks are highlighted in the close-up at the maximum vertical displacement reported in Fig. 9, when heavy vehicles enter and exit the bridge span and their wheels hit the bridge expansion joints at the ends of each span, at the start and the end of the vehicle passage, i.e., start and end of the vertical deflections of Targets T1 and T2. This apparent displacement is consequently a noise also present in T2 and T1.

The power spectrum of the pixel displacement history (Fig. 10) shows the main peak associated with the first mode of the deck (frequency 3.52 Hz) that is captured by both T1 and T2 targets. Differently, the power spectrum of T3 is not affected by the vibration of the deck, as expected, given the very high vertical stiffness of the pier [Fig. 6(a)]. It is also observed that in the 20–25 Hz range, the three signals T1, T2, and T3 have the same spectrum. This can be explained by recalling that the vertical and rotational motion of the camera induces the same apparent pixel displacement at each target; accordingly, this range of frequencies is imputable to the vibration of the video camera due to the impulses created by the traffic on the bridge and filtered by the tripod–camera system.

Vertical Displacements with Camera Movement Compensation

By subtracting the pixel fluctuations of Reference target T3 from the pixel displacements of the other two targets, the effects of the

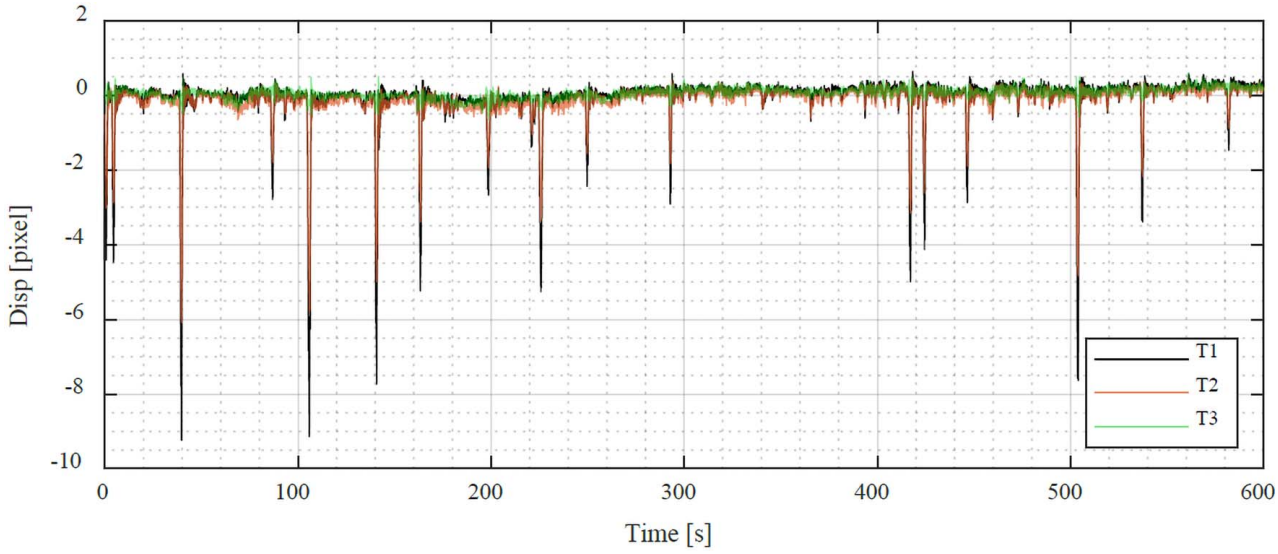

Fig. 8. Vertical (Y) displacements in pixels of Targets T1, T2, and T3 for the 10-min window.

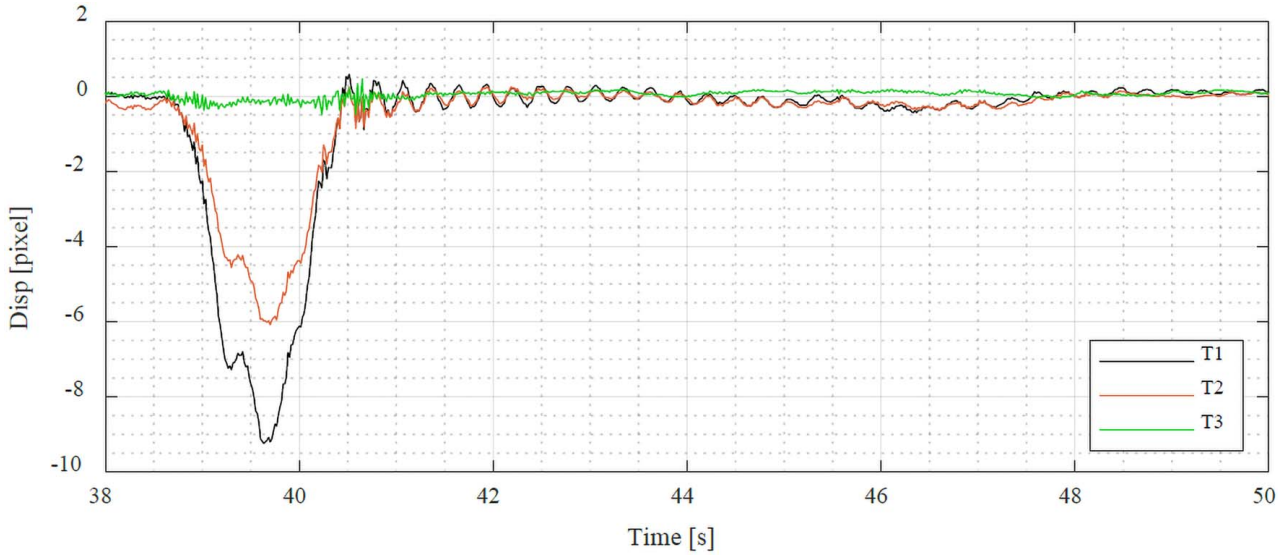

Fig. 9. Vertical (Y) displacements in pixels of Targets T1, T2, and T3, close-up at maximum displacement.

in-plane translation and pitch rotation of the sensor are removed. The obtained pixel displacements are then converted into physical displacements through the scale factor ratio R_{SF}. The result of the correction is shown in Fig. 11, where the corrected (T1CORR) and uncorrected (T1) vertical deflections of the target at the mid-span are compared to the vertical deflections measured by the DT in the same position. Fig. 12 shows a close-up around the maximum displacement (between 38 and 50 s) where the effect on the camera vibration is even more evident: the corrected signal T1CORR is closer to the measure provided by the contact sensor (DT), and an important component of video camera vibration, highlighted at the end of the truck passage (around 40 s), is eliminated. Looking at the last portion in the considered 10-min window (Fig. 13), the effect of the correction is even clearer, especially regarding the baseline, influenced by the microsettlements and movements of the camera tripod accumulated during the 10-min acquisition window. It is also interesting to see that in the 580–594 s time interval, at least four vehicles crossed the monitored bridge span (580, 582, 586, and 593 s), producing perturbations in the video camera; however, the corrected results are still in good agreement with the transducer, even if displacements are small (lower than 1 mm).

The comparison between corrected (T1CORR) and uncorrected (T1) video measures in terms of power spectra (Fig. 14) reveals that the correction based on the assumed stationary target on the pier cap produces a significant reduction of the spectral noise in the 17–27 Hz range [identified as the range of camera vibrations in the experimental campaign described in Micozzi et al. (2023, 2024), where two accelerometers were installed on the tripod on the same plate of the video camera] and a negligible increment in the 4–12 Hz range.

The comparison in Fig. 14 between the image-derived measurement spectra, the spectrum from the transducer displacement, and the spectrum from the accelerometer also shows that the video camera, regardless of the implemented correction, provides a very good estimate of the frequency of the first mode. The computer vision system is, however, unable to capture the superior modes of vibration of this bridge (the frequency of the second mode is 10.18 Hz)

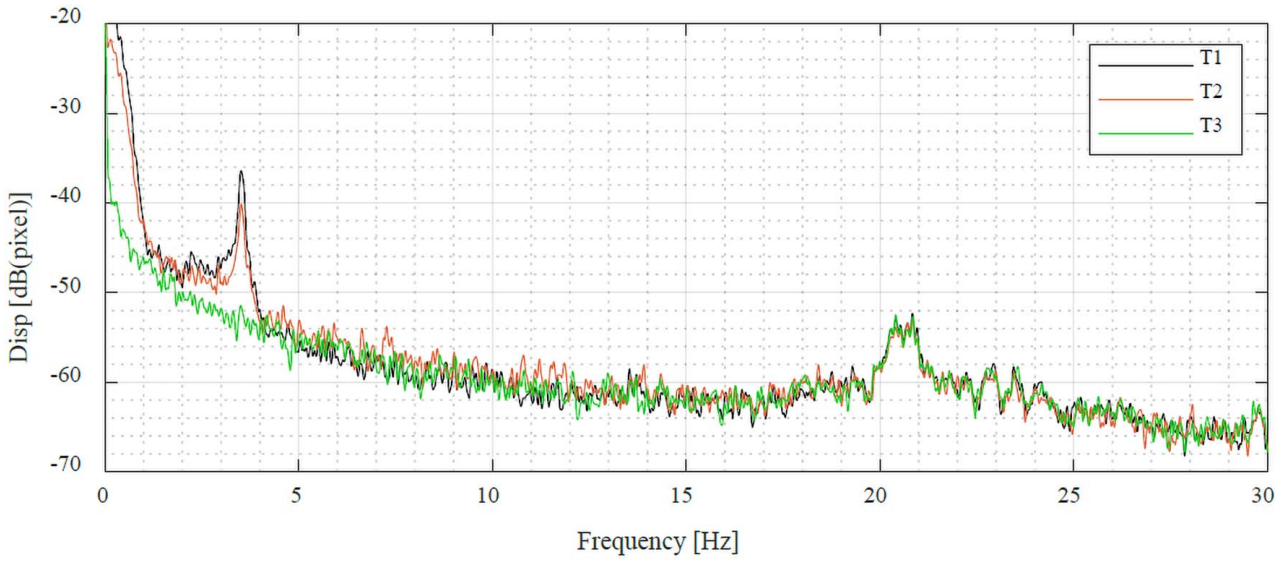

Fig. 10. Power spectrum of the vertical (Y) displacements in pixel of Targets T1, T2, and T3 from the 10-min window.

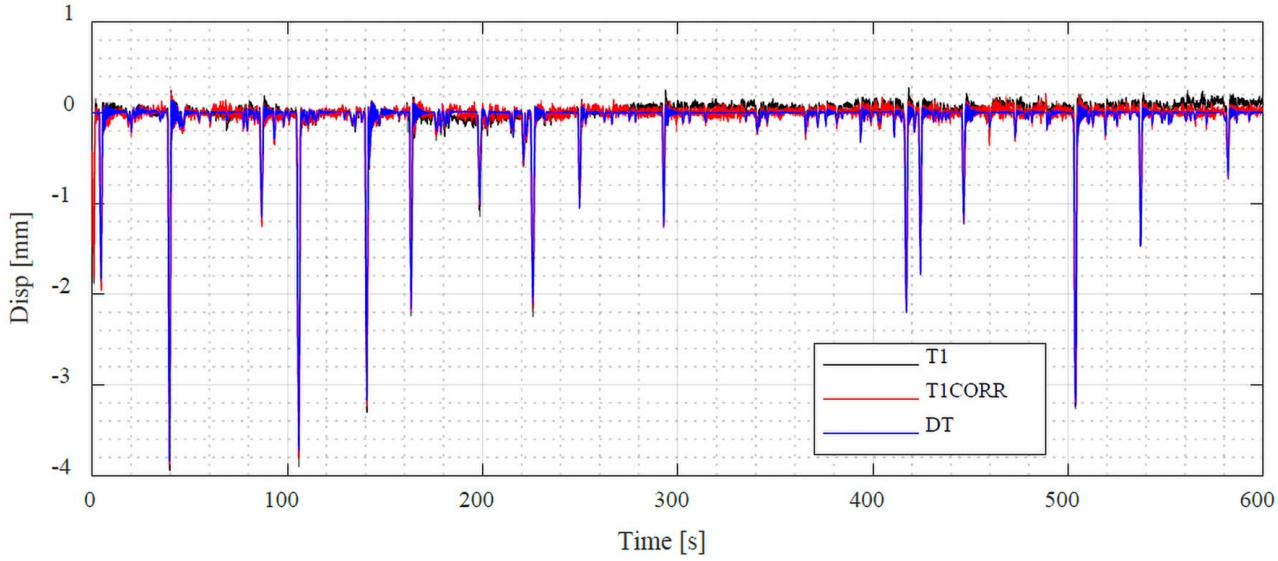

Fig. 11. Vertical (Y) displacements of Target T1 as acquired and corrected (T1CORR), compared with the DT from the 10-min window.

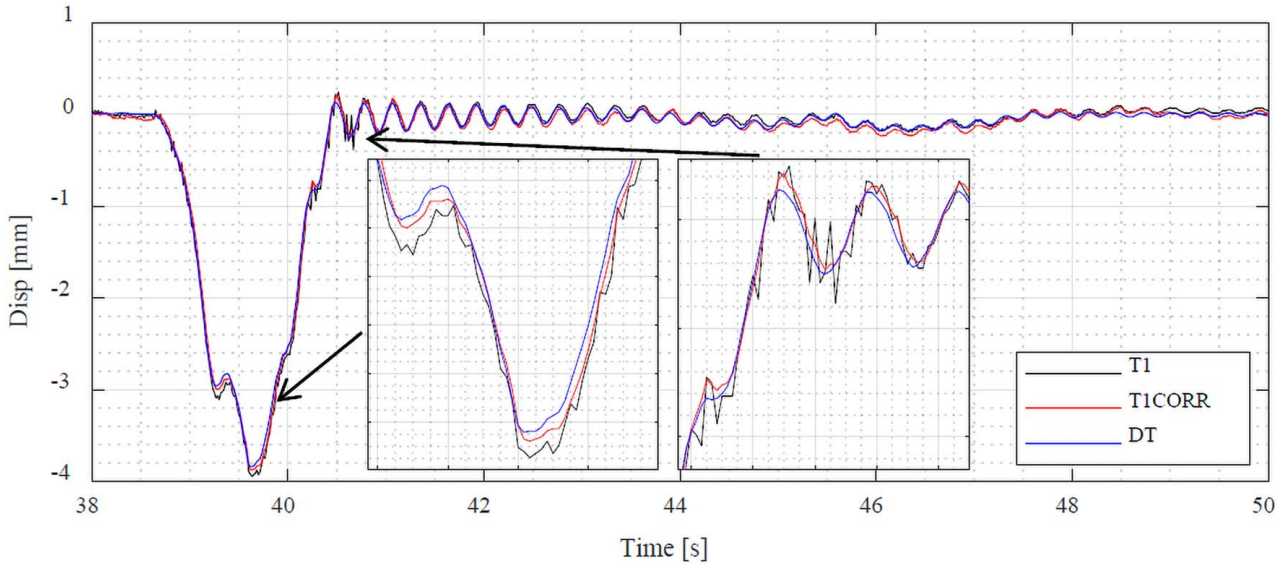

Fig. 12. Vertical (Y) displacements of Target T1 as acquired and corrected (T1CORR), compared with the DT, close-up at maximum displacement.

due to the noise floor that covers the frequency peaks after the first one and the decay of the contribution of the superior modes to the displacements. Such limitation is shown in this study to clearly highlight the possibilities currently offered by the presented computer vision solution, recalling that the identification of the higher modes of vibration is not the objective of the presented methodology, focused on the direct measurement of displacements.

Quantitative Analysis of the Measurement Accuracy

To provide a quantitative comparison between the accuracy of the measured displacement with and without the correction based on the stationary target, Table 1 reports the highest 15 peaks for T1, T1CORR, and DT. The error on the peaks is already small for T1, with a mean of 0.039 mm and a standard deviation of 0.077 mm. However, these values are even lower after applying the correction; in this case, the mean error decreases to 0.011 mm and the standard deviation decreases to 0.052 mm. Other metrics were used to compare the measurements during the entire 10-min acquisition window and not considering only the peaks; they are the RMS error (RMSE), the weighted RMS error (WRMSE), the coefficient of determination (RSQ), and the time response assurance criterion (TRAC), respectively

$$\text{RMSE} = \sqrt{\sum_{i=1}^{N} \frac{(d_{\text{VC},i} - d_{\text{DT},i})^2}{N}} \tag{4}$$

$$\text{WRMSE} = \sqrt{\frac{\sum_{i=1}^{N} \frac{|d_{\text{VC},i}|(d_{\text{VC},i} - d_{\text{DT},i})^2}{N}}{\sum_{i=1}^{N} d_{\text{VC},i}}} \tag{5}$$

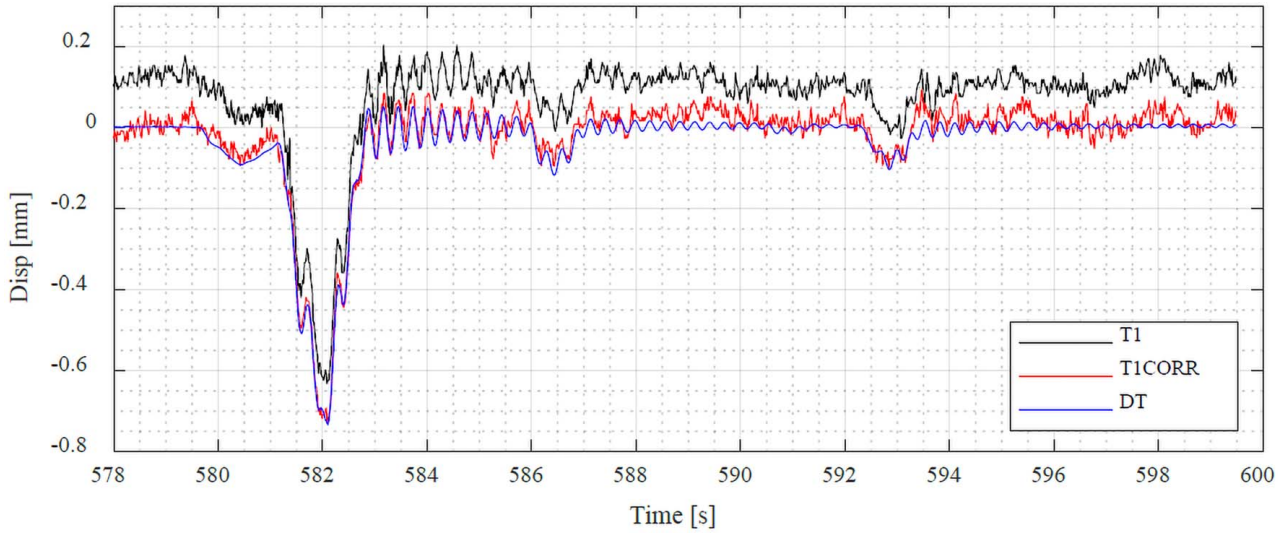

Fig. 13. Vertical (Y) displacement (mm) of Target T1 as acquired and corrected (T1CORR) and the displacement measured by the transducer (DT), close-up of the last 22 s of the 10-min window.

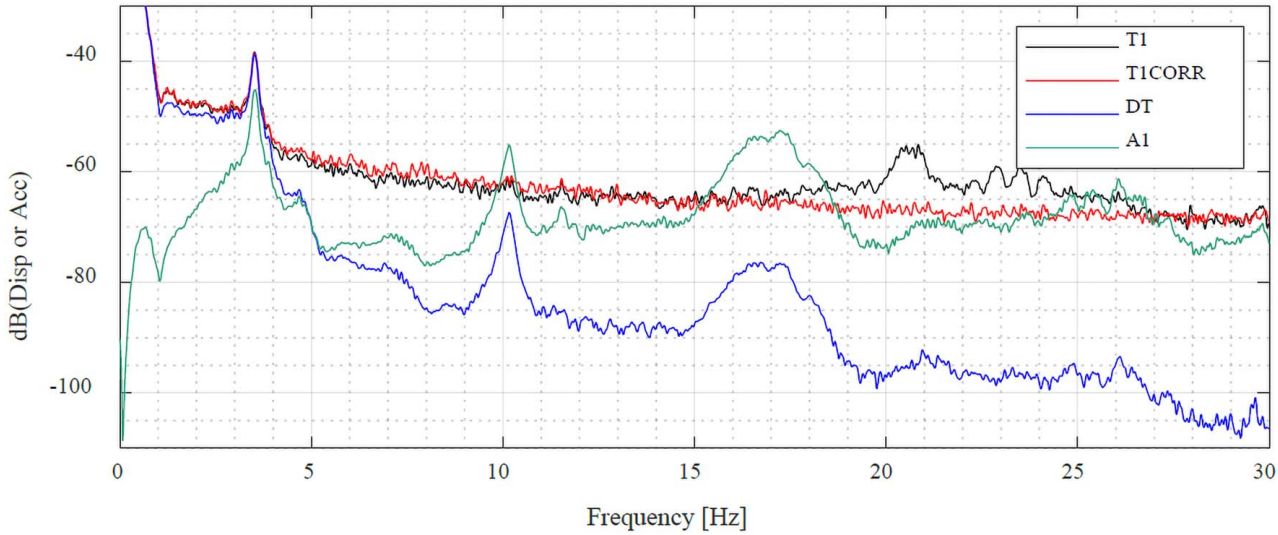

Fig. 14. Power spectra comparison of the Y displacement (mm) of targets T1 as acquired and corrected (T1CORR), the displacement measured by the transducer (DT), and the acceleration measured by the accelerometer (A1) for the 10-min window.

$$\mathrm{RSQ} = 1 - \frac{\sum_{i=1}^{N} (d_{\mathrm{VC},i} - d_{\mathrm{DT},i})^2}{\sum_{i=1}^{N} (d_{\mathrm{VC},i} - \mu)^2} \tag{6}$$

$$\mathrm{TRAC} = \frac{\left(\sum_{i=1}^{N} d_{\mathrm{VC},i} \cdot d_{\mathrm{DT},i}\right)^2}{\left(\sum_{i=1}^{N} d_{\mathrm{VC},i}^2\right) \cdot \left(\sum_{i=1}^{N} d_{\mathrm{DT},i}^2\right)} \tag{7}$$

where $d_{\mathrm{VC},i}$ and $d_{\mathrm{DT},i} = i$th readings from the video camera and displacement transducer, respectively; $N =$ total number of readings in the considered signal; and $\mu =$ mean value of the N readings $d_{\mathrm{VC},i}$. The computed RMSE on the whole signal is 0.067 mm for T1 and 0.036 mm for T1CORR, while the WRMSE, with the weight vector equal to the displacement DT, is 0.084 mm for T1 and 0.048 mm for T1CORR. The RMSE values suggest a very low difference between the camera results and the reference measures. Moreover, the WRMSE confirms, on the whole signal, the RMSE values computed only on the peaks (see the last row of Table 1). The quality of the results is confirmed by the value of TRAC, which is a metric related only to the shape of the signals, disregarding the amplitude; i.e., it is not affected by R_{SF}. The TRAC value is equal to 0.958 for T1 and 0.987 for T1CORR. The RSQ values, equal to the TRAC values, confirm the accuracy of the automatically determined R_{SF} because any error introduced by the scale factor can only reduce the RSQ metric. The effect of the correction using the reference target is always evident, confirming that the compensation of the measurements has a beneficial effect on the precision of the measurements, both for the condition of the loaded bridge (correction of the camera vibration induced by the traffic) and unloaded bridge (correction of the deviation of the camera due to tripod settlements and their accumulations).

Horizontal Transverse Displacements

A single video camera in the configuration proposed in this work, with the optical axis parallel to the bridge axis, allows for directly

Table 1. Values, differences, and statistical parameters of the highest 15 peaks of the 10-min window

Time (s)	T1 (mm)	T1COR (mm)	DT (mm)	ΔT1 (mm)	ΔT1COR (mm)	Err % T1	Err % T1COR
4.6	−1.91	−1.96	−1.83	0.080	0.122	−4.3	−6.7
39.6	−3.92	−3.87	−3.84	0.080	0.029	−2.1	−0.8
86.8	−1.15	−1.20	−1.15	−0.002	0.050	0.1	−4.3
105.9	−3.90	−3.80	−3.72	0.179	0.077	−4.8	−2.1
140.8	−3.26	−3.22	−3.17	0.091	0.048	−2.9	−1.5
163.4	−2.23	−2.18	−2.16	0.068	0.017	−3.2	−0.8
198.6	−1.14	−0.92	−1.00	0.142	−0.080	−14.1	8.0
225.9	−2.23	−2.13	−2.10	0.132	0.030	−6.3	−1.4
249.7	−1.04	−1.03	−1.06	−0.014	−0.023	1.4	2.2
292.8	−1.20	−1.24	−1.25	−0.049	−0.015	3.9	1.2
417.2	−2.14	−2.20	−2.20	−0.069	−0.009	3.1	0.4
424.5	−1.76	−1.79	−1.79	−0.025	0.000	1.4	0.0
446.6	−1.20	−1.21	−1.21	−0.004	0.004	0.3	−0.4
504.0	−3.21	−3.18	−3.20	0.008	−0.026	−0.3	0.8
537.1	−1.44	−1.41	−1.47	−0.039	−0.064	2.6	4.4
—	—	—	Mean	0.039	0.011	−1.7	−0.1
—	—	—	STD	0.077	0.052	4.6	3.4
—	—	—	RMSE	0.084	0.052	—	—

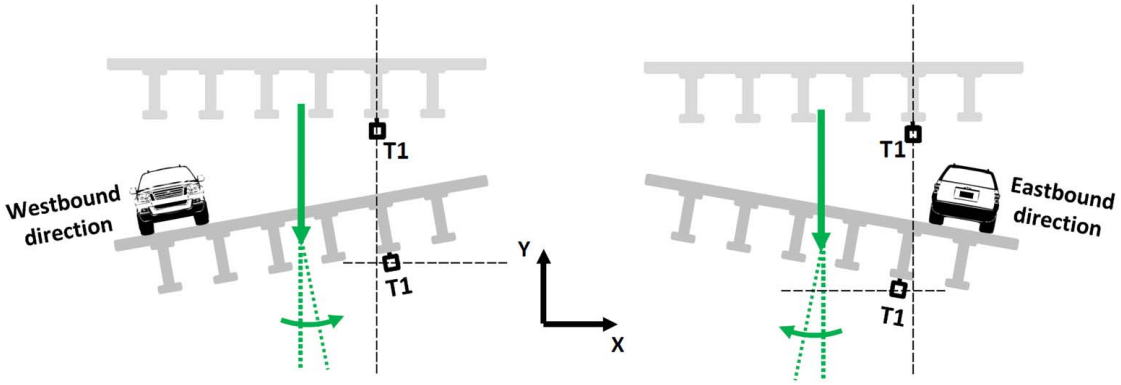

Fig. 15. Identification of the direction of the vehicle from horizontal and vertical displacements.

measuring the horizontal transverse displacements in addition to the vertical displacements. The horizontal transverse displacements are a consequence of the rotation of the cross section of the bridge deck (Fig. 15), and their sign defines whether a vehicle is in the lane directly above the monitored beam (second beam from the right) or on the other side. A negative horizontal displacement indicates a clockwise rotation of the deck, as in the case of a vehicle driving Eastbound, while a positive horizontal displacement indicates a

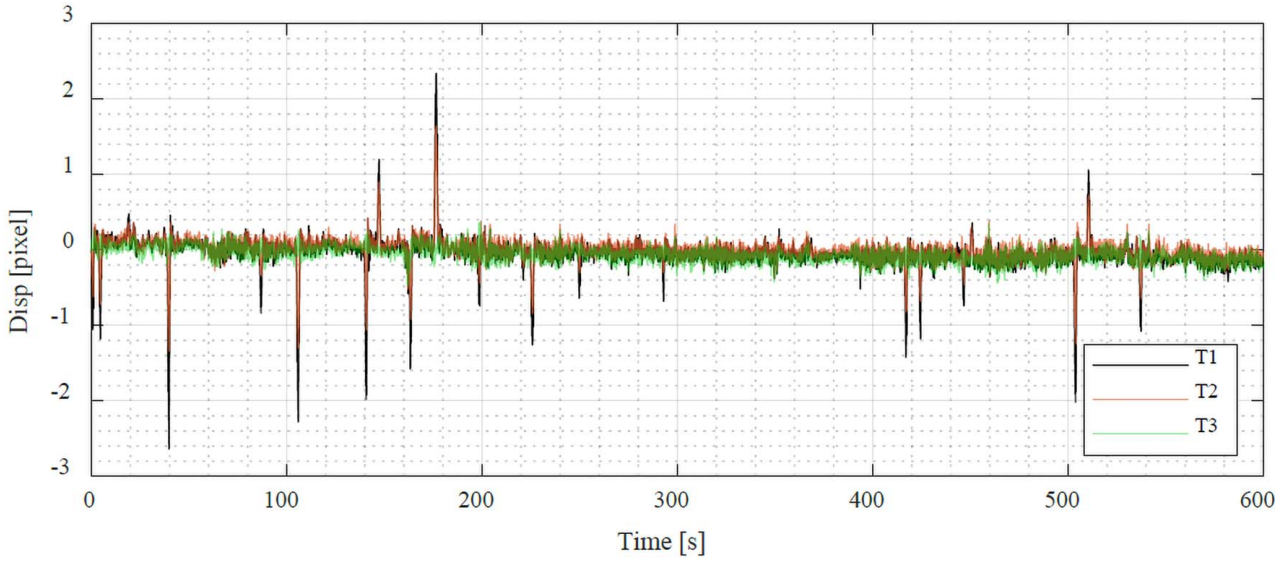

Fig. 16. Horizontal (X) transverse displacements in pixels of Targets T1, T2, and T3 for the 10-min window.

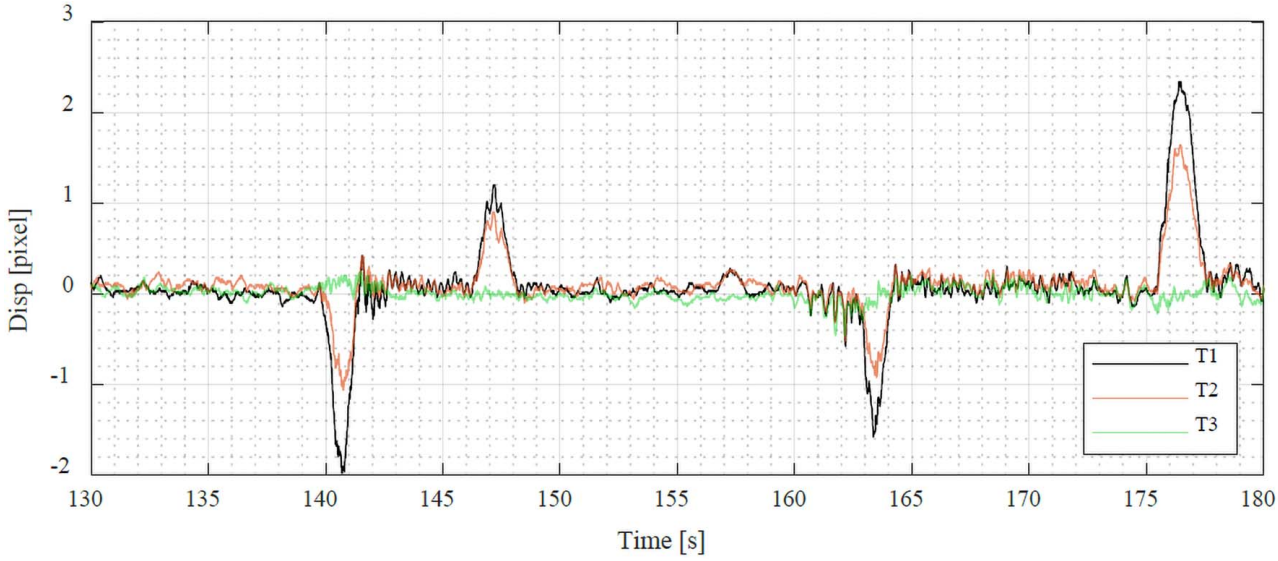

Fig. 17. Horizontal (X) transverse displacements in pixels of Targets T1, T2, and T3, close-up at peaks with different sign.

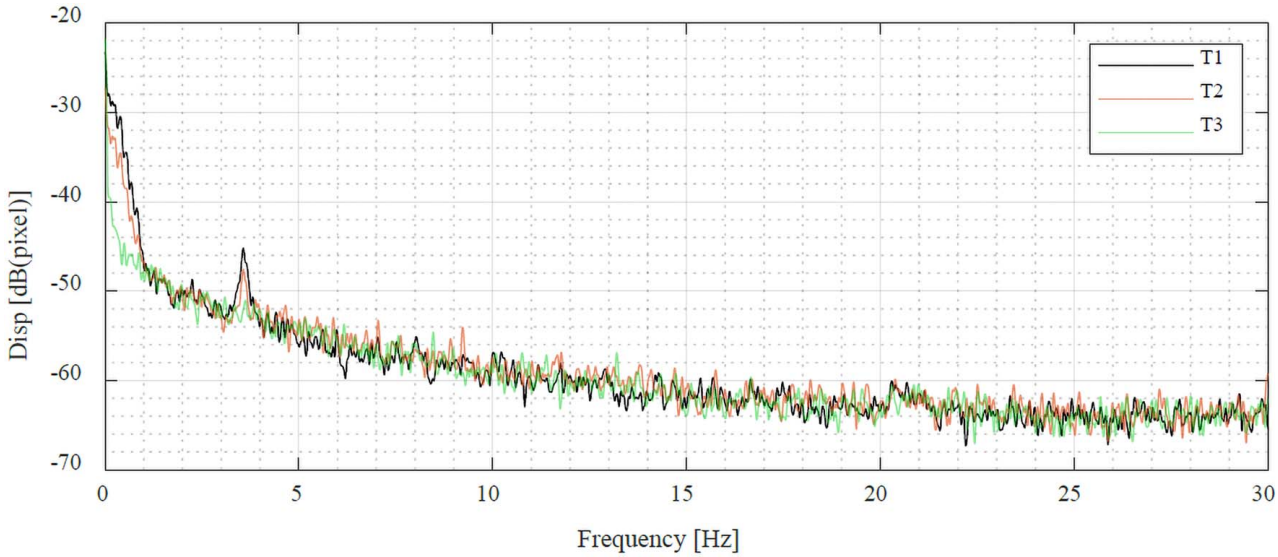

Fig. 18. Power spectrum of the horizontal (X) transverse displacements in pixels of Targets T1, T2, and T3 from the 10-min window.

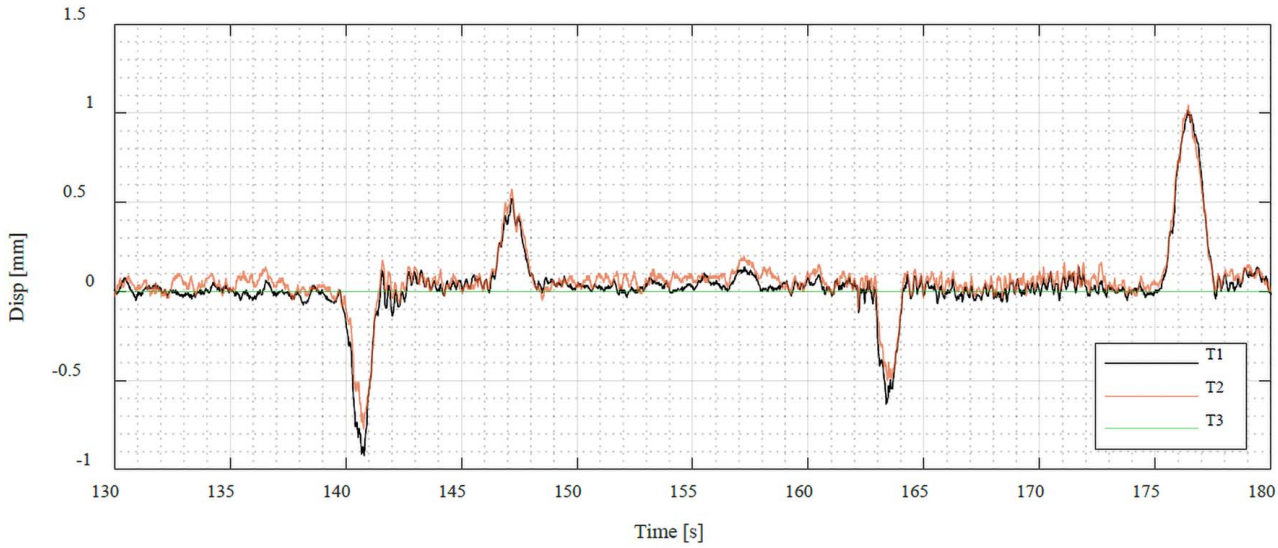

Fig. 19. Horizontal (X) transverse displacements of Target T1 as acquired and corrected (T1CORR), close-up at peaks with different signs.

counterclockwise rotation, as in the case of a vehicle driving Westbound.

Fig. 16 shows the displacements in pixels of the three targets, before the conversion of the measurements into physical dimensions through the scale factor. As expected, the horizontal displacements are notably smaller than the vertical displacements. Fig. 17 provides a close-up between 130 and 180 s where four peaks, two positives, and two negatives, are found. The noise in the readings of Target T3, assumed as a fixed point, is easily observed, although its relations with the transiting vehicles appear less pronounced. It is also observed that the power spectra shown in Fig. 18 are regularly decaying after the only clearly identifiable peak corresponding to the first mode of the bridge, as already commented for Fig. 14. The corrected and uncorrected horizontal transverse displacements for Target T1 are shown in Fig. 19 for the same close-up used in Fig. 17. Once again, the beneficial effect of the correction is highlighted.

The vertical (negative) displacement peaks are always aligned with the negative horizontal transverse displacement peaks (clockwise rotations), as observed in Fig. 20, confirming that the most significant effects on the monitored beam (second from the right) are a consequence of the transit of a vehicle directly above the target (Eastbound direction). Conversely, positive horizontal displacements (counterclockwise rotations) are aligned to smaller vertical displacements, suggesting the passage of a vehicle on the other side of the bridge (Westbound direction) with respect to the monitored beam. For example, Fig. 21 highlights the passage of four heavy vehicles, two on the lane monitored (141 and 163 s) and two on the opposite side (147 and 176 s).

Influence of Target Blurring

As already mentioned, in the proposed configuration, each deck cross section has a different distance from the video camera.

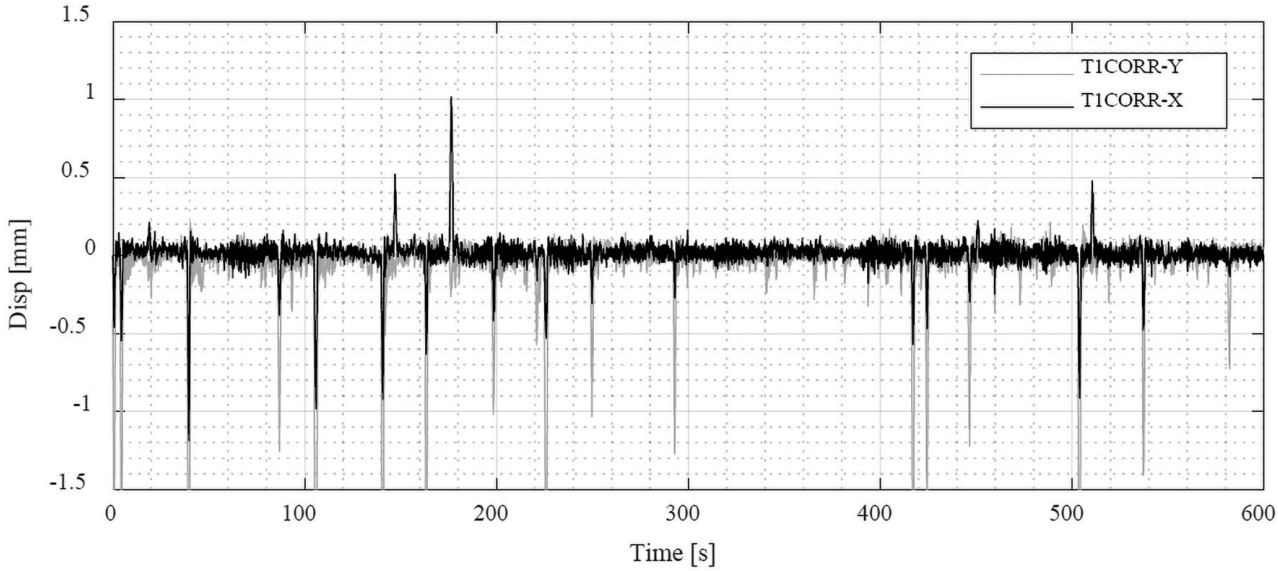

Fig. 20. Vertical (Y) and horizontal (X) transverse displacements (corrected with respect to the stationary target) for the 10-min window.

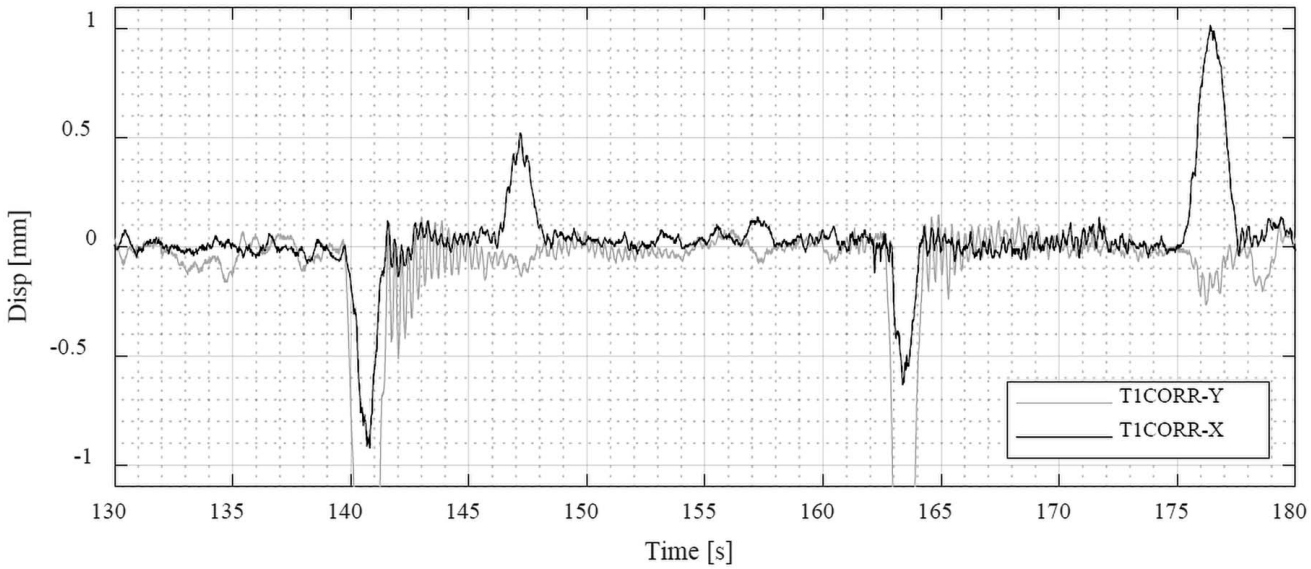

Fig. 21. Vertical (Y) and horizontal (X) transverse displacements (corrected with respect to the stationary target), close-up at peaks with different signs.

Accordingly, only one target can be precisely focused, while all other targets will inevitably be out of focus. The influence of target blurring caused by targets that are partially out of focus was verified in the considered case study, comparing the displacements extracted in real time from the targets as seen by the video camera to the displacement extracted from the recorded backup video footage after the application of a Gauss filter that adds artificial blurring, which simulates the increased out of focus. The original view of the targets and their blurred counterparts are reported in Fig. 22, while the extracted vertical displacements are compared in Fig. 23. Marginal differences can be observed between the two measurements. Hence, the robustness of the proposed hardware configuration (video camera below the bridge deck with optical axis parallel to the bridge axis, focusing only on one of the multiple cross sections that can be monitored) with respect to target blurring as caused by partial out of focus is confirmed when used together with an area-based algorithm for displacement extraction. It should be remarked that other factors may also cause blurring of the acquired images, such as sensor noise that typically increases with increasing sensor gain, motion blur when the adopted FPS and exposition time are not adequate to the speed of the object whose displacement is being tracked, and visual disturbances generated by uneven illumination conditions. In the presented application, the sensor gain was kept to low, the adopted video camera exhibited minimal temporal dark noise, motion blur was extremely limited because the targets moved only by a small fraction of a pixel between subsequent frames (60 FPS and at 0.005 s exposure time), the targets were relatively small, and no uneven illumination conditions on their surfaces were observed.

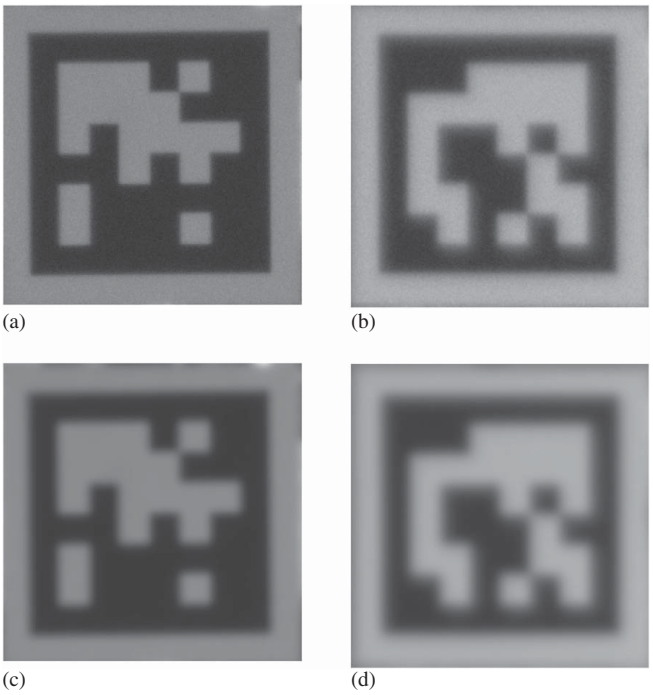

Fig. 22. Original views and blurred views of Targets T1 and T3: (a) Target T1 as seen by the camera; (b) Target T3 as seen by the camera; (c) Target T1 with additional blurring; and (d) Target T3 with additional blurring.

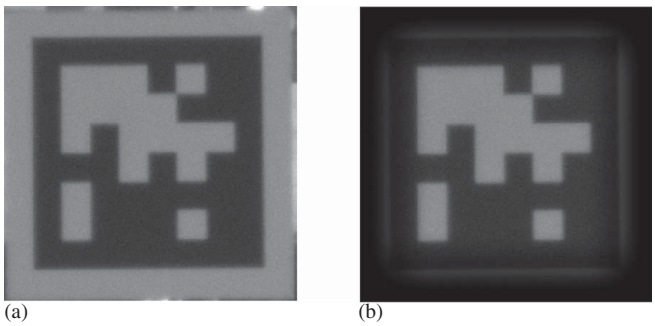

Fig. 24. Visualization of the modifications introduced by windowing: (a) before windowing; and (b) after windowing.

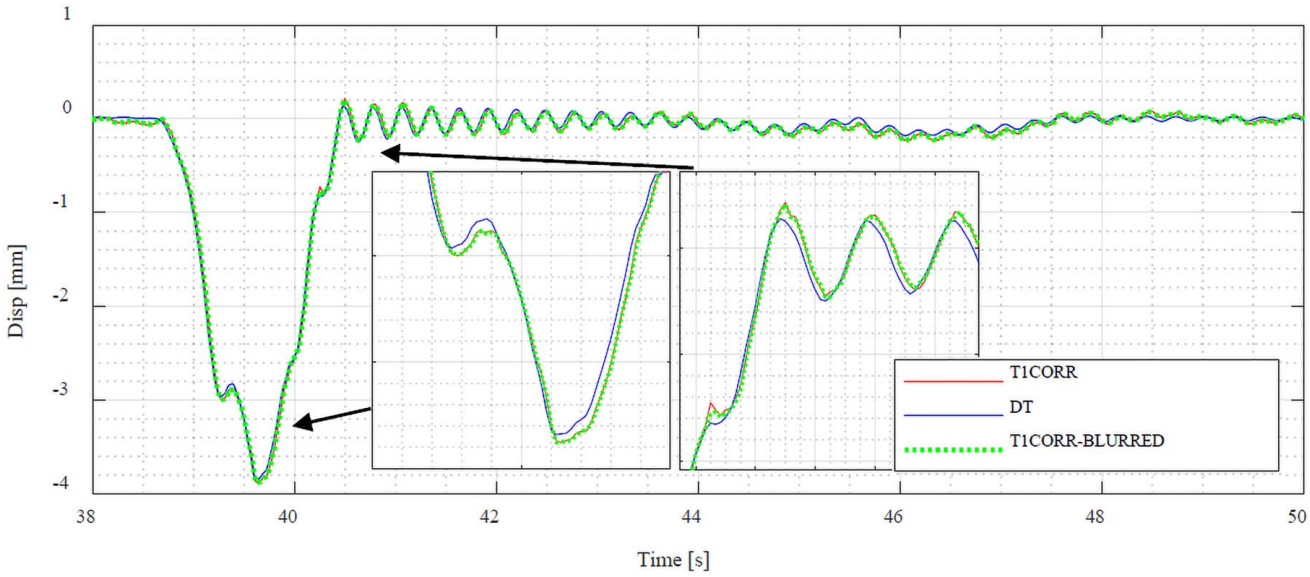

Fig. 23. Vertical (Y) displacements of Target T1 corrected (T1CORR) from the original views and (T1CORR-BLURRED) from the blurred views of Targets T1 and T3, comparison with the DT, close-up at maximum displacement.

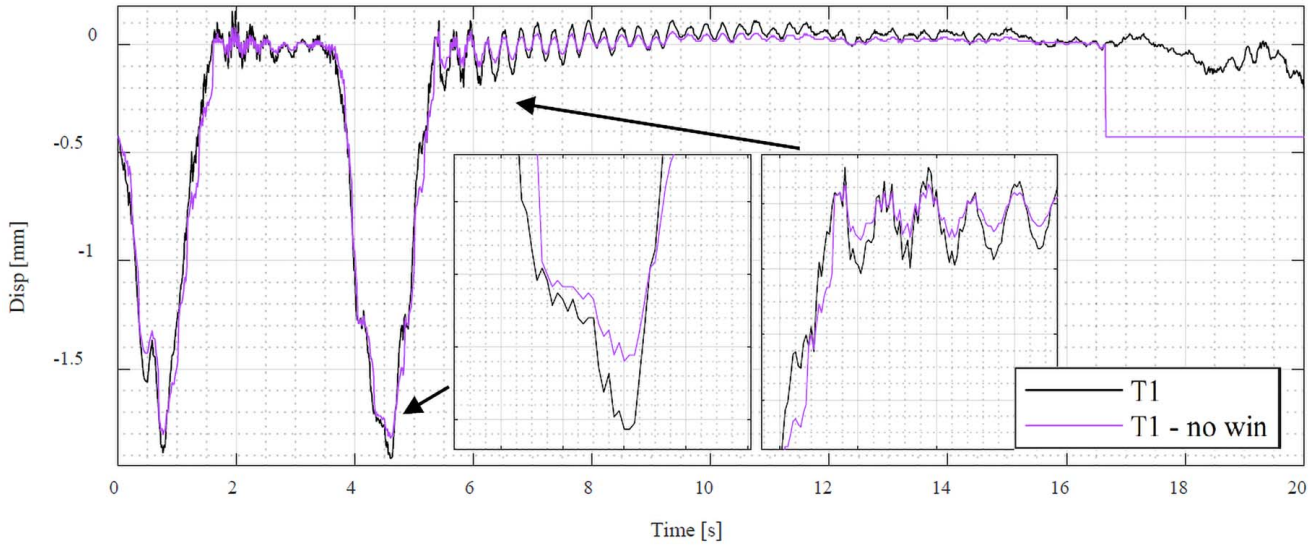

Fig. 25. Example of inaccurate and unstable displacement tracking when no windowing is used in a small ROI (sub-ROI of Target T1).

Influence of Subdomain Windowing

The adopted algorithm for displacement tracking includes a windowing processing for each target subdomain, implemented, as already illustrated, to reduce noise and increase the robustness of displacement measurements, especially for small sub-ROIs, as is the case of the proposed multitarget workflow. Fig. 24 provides a visualization of the effect of windowing, showing Target T1 as it appears before and after windowing. Fig. 25 compares the displacements obtained in real time with the combined use of the template matching algorithm and windowing with the displacement obtained from the recorded backup video footage using the template matching algorithm without windowing. The differences between the two time histories appear small during the first 17 s; afterward, the algorithm operating on the target without windowing fails, unable to extract displacement variations from one frame to the following one, leaving erroneously fixed the position of the monitored target. This is an example of the instability problems that might be encountered when images of small size are processed without windowing. Theoretically, such problems could be avoided using larger ROIs, without necessarily relying on windowing. However, in this way, the approach would be computationally inefficient compared to the proposed workflow and would not permit attaining the same real-time capacity in terms of displacement components that are extracted at a given acquisition frequency compatible with the dynamic monitoring of displacements of bridge decks.

Conclusions

This study presented a computer vision workflow for real-time dynamic monitoring of the vertical and horizontal transverse displacements of the cross sections of bridge decks. The proposed original contributions are as follows:

1. The hardware configuration (video cameras placed under the bridge deck, near one abutment, or an intermediate pier, with its optical axis parallel to the bridge axis) specifically oriented to bridge deck monitoring differs from the commonly adopted installations, both in terms of displacements that can be measured (bending and torsion can be identified with a single video camera) and in terms of installation replicability for most bridge deck geometries.
2. For the first time in structural monitoring, an area-based upsampled cross-correlation FFT template matching algorithm for displacement tracking (recognized for its performance and efficiency) was combined with a windowing function (known for avoiding problems in FFT matching when handling small images) to achieve significant improvements in noise reduction and measurement repeatability, while working synergistically way with the proposed hardware configuration. It also shows low sensitivity to targets that are not precisely focused, an inevitable situation given the varying distances of bridge cross sections from the video camera.
3. Image subdomains continually recentered on each monitored target were introduced; the displacement-tracking algorithm searches for the new position of the target within this subdomain, with significant benefits in terms of computational time, allowing real-time tracking of the movements of multiple targets acquired by the same video camera at sampling frequencies sufficient to correctly capture the dynamic displacement response of the bridge deck.
4. Use of coded targets allowed automated camera calibration (identification of the scale factor for each target), automated definition of the initial position of sub-ROIs, and automated compensation for the movements of the video camera by referencing a stationary target, which is always available in the proposed configuration.

The application to a real case study consisting of a post-tensioned concrete bridge having a 32-m-long simply supported span under vehicular traffic provided the following outcomes:

1. The adopted combination of computer vision software and hardware gave very accurate measurements, reaching submillimetric equivalence with a displacement transducer under dynamic operational conditions, both in terms of peak measures and for the entire signal.
2. The vision-based measurements of the displacements permitted identification of the first vibration frequency of the bridge deck, matching the frequency derived from the displacement transducer and the accelerometer. The noise for frequencies above 10 Hz precluded the possibility of identifying the subsequent vibration modes from the video camera. Indeed, the considered

bridge appears as a severe test for image-based monitoring because of the small amplitude of displacements and the higher modal frequency range compared to other, more deformable bridges studied previously. Currently, if the goal of deck monitoring is its complete frequency analysis that extends to the higher modes, accelerometers are confirmed as the most suitable instrumentation.

3. Vision-based monitoring, taking advantage of optimized image-processing algorithms, automated target identification for camera calibration, and camera disturbance compensation, appears to be an unrivalled solution for comprehensive monitoring of the displacements of a bridge deck, both in quasi-static and dynamic conditions, given that no other technologies can deliver the same information with a single cost-effective video camera.
4. Although tested for a short-span post-tensioned concrete bridge, the presented workflow has potential applications to other bridge typologies and longer spans, under the condition that the video cameras and coded targets can be installed according to the proposed configuration. It is noted that deflection is a cubic function of the span length, while the precision of computer vision measurements ideally reduces linearly with the distance. Although humidity and heat haze can make things more complicated, they are expected to have a limited influence on the configuration proposed because the field of view is in a constantly shaded volume of air. Nevertheless, experimental full-scale applications are required to provide insight into these aspects in the case of longer bridge spans.
5. The presented workflow opens the way toward image-based long-term monitoring, thanks to its real-time capabilities (no need to store large video files) and the convenient configuration (complete description of the transverse movements of the deck in multiple cross sections, a stationary point for compensating disturbance in the video camera expected to be always included in the camera field of view, protection from direct sunlight and atmospheric events, and no interferences with bridge normal operational conditions). However, more studies are needed to evaluate the long-term performance of the video cameras and the influence of varying environmental and illumination conditions, both at daytime and nighttime, including strategies for their automated compensation.

Data Availability Statement

Data recorded from the measurements used in this study are available from the corresponding author upon reasonable request.

Acknowledgments

This research received economic support from the FABRE Consortium and the Italian Ministry of University and Research project "TIMING: Time evolution laws for improving the structural reliability evaluation of existing post-tensioned concrete deck bridges" (Grant No. P20223Y947).

Author Contributions

Fabio Micozzi: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing—original draft, Writing—review and editing; Michele Morici: Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing—original draft; Alessandro Zona: Conceptualization, Formal analysis, Investigation, Methodology, Resources, Software, Supervision, Writing—original draft, Writing—review and editing; Andrea Dall'Asta: Methodology, Project administration, Resources, Supervision, Writing—review and editing.

References

Aliansyah, Z., K. Shimasaki, T. Senoo, I. Ishii, and S. Umemoto. 2021. "Single-camera-based bridge structural displacement measurement with traffic counting." *Sensors* 21 (13): 4517. https://doi.org/10.3390/s21134517.

Alipour, M., S. J. Washlesky, and D. K. Harris. 2019. "Field deployment and laboratory evaluation of 2D digital image correlation for deflection sensing in complex environments." *J. Bridge Eng.* 24 (4): 04019010. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001363.

Bunce, A., D. Hester, S. Taylor, J. Brownjohn, F. Huseynov, and Y. Xu. 2023. "A robust approach to calculating bridge displacements from unfiltered accelerations for highway and railway bridges." *Mech. Syst. Sig. Process.* 200 (1): 110554. https://doi.org/10.1016/j.ymssp.2023.110554.

Catbas, F. N., R. Zaurin, M. Gul, and H. B. Gokce. 2012. "Sensor networks, computer imaging, and unit influence lines for structural health monitoring: Case study for bridge load rating." *J. Bridge Eng.* 17 (4): 662–670. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000288.

Chen, H. P. 2018. *Structural health monitoring of large civil engineering structures*. New York: Wiley Blackwell.

Debees, M., F. Luleci, C.-Z. Dong, and F. N. Catbas. 2024. "Investigating the increase in load rating and reliability of a prestressed concrete bridge when utilizing field-derived distribution and impact factors." *J. Bridge Eng.* 29 (3): 04023123. https://doi.org/10.1061/JBENF2.BEENG-6485.

Dhanasekar, M., P. Prasad, J. Dorji, and T. Zahra. 2019. "Serviceability assessment of masonry arch bridges using digital image correlation." *J. Bridge Eng.* 24 (2): 04018120. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001341.

Dong, C.-Z., and F. N. Catbas. 2021. "A review of computer vision–based structural health monitoring at local and global levels." *Struct. Health Monit.* 20 (2): 692–743. https://doi.org/10.1177/1475921720935585.

Farrar, C. R., and K. Worden. 2013. *Structural health monitoring: A machine learning perspective*. New York: Wiley.

Feng, D., and M. Q. Feng. 2017. "Experimental validation of cost-effective vision-based structural health monitoring." *Mech. Syst. Sig. Process.* 88 (1): 199–211. https://doi.org/10.1016/j.ymssp.2016.11.021.

Feng, D., and M. Q. Feng. 2018. "Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection—A review." *Eng. Struct.* 156 (1): 105–117. https://doi.org/10.1016/j.engstruct.2017.11.018.

Feng, D., and M. Q. Feng. 2021. *Computer vision for structural dynamics and health monitoring*. 1st ed. Hoboken, NJ: Wiley.

Ghyabi, M., L. C. Timber, G. Jahangiri, D. Lattanzi, H. W., Shenton III, M. J. Chajes, and M. H. Head. 2023. "Vision-based measurements to quantify bridge deformations." *J. Bridge Eng.* 28 (1): 05022010. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001960.

Gindy, M., R. Vaccaro, H. Nassif, and J. Velde. 2008. "A state-space approach for deriving bridge displacement from acceleration." *Comput.-Aided Civ. Infrastruct. Eng.* 23 (4): 281–290. https://doi.org/10.1111/j.1467-8667.2007.00536.x.

Guizar-Sicairos, M., S. T. Thurman, and J. R. Fienup. 2008. "Efficient subpixel image registration algorithms." *Opt. Lett.* 33 (2): 156–158. https://doi.org/10.1364/OL.33.000156.

Hou, X., X. Yang, and Q. Huang. 2005. "Using inclinometers to measure bridge deflection." *J. Bridge Eng.* 5 (564): 564–569. https://doi.org/10.1061/(ASCE)1084-0702(2005)10:5(564).

Krogius, M., A. Haggenmiller, and E. Olson. 2019. "Flexible layouts for fiducial tags." In *Proc., IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 1898–1903. London, UK: IEEE.

Luo, K., X. Kong, J. Zhang, J. Hu, J. Li, and H. Tang. 2023. "Computer vision-based bridge inspection and monitoring: A review." *Sensors* 23 (18): 7863. https://doi.org/10.3390/s23187863.
MathWorks. 2024. *MATLAB version: 23.2.0 (R2023b)*. Natick, MA: MathWorks.
Micozzi, F., M. Morici, A. Zona, and A. Dall'Asta. 2023. "Vision-based structural monitoring: Application to a medium-span post-tensioned concrete bridge under vehicular traffic." *Infrastructures* 8 (10): 152. https://doi.org/10.3390/infrastructures8100152.
Micozzi, F., M. Morici, A. Zona, and A. Dall'Asta. 2024. "Vision-based dynamic monitoring of a post-tensioned concrete bridge under vehicular traffic." *Procedia Struct. Integrity* 62: 848–855. https://doi.org/10.1016/j.prostr.2024.09.114.
Moreu, F., J. Li, H. Jo, R. E. Kim, S. Scola, B. F. Spencer, and J. M. LaFave. 2016. "Reference-free displacements for condition assessment of timber railroad bridges." *J. Bridge Eng.* 21 (2): 04015052. https://doi.org/10.1061/(asce)be.1943-5592.0000805.
Narasimhan, S., and Y. Wang. 2020. "Noncontact sensing technologies for bridge structural health assessment." *J. Bridge Eng.* 25 (6): 02020001. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001560.
Olson, E. 2011. "Apriltag: A robust and flexible visual fiducial system." In *Proc., IEEE Int. Conf. on Robotics and Automation*, 3400–3407. London, UK: IEEE.
Ozdagli, A. I., J. A. Gomez, and F. Moreu. 2017. "Real-time reference-free displacement of railroad bridges during train-crossing events." *J. Bridge Eng.* 22 (10): 04017073. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001113.
Park, K.-T., S.-H. Kim, H.-S. Park, and K.-W. Lee. 2005. "The determination of bridge displacement using measured acceleration." *Eng. Struct.* 27 (3): 371–378. https://doi.org/10.1016/j.engstruct.2004.10.013.
Scozzese, F., and A. Dall'Asta. 2024. "Nonlinear response characterization of post-tensioned R.C. bridges through Hilbert–Huang transform analysis." *Struct Cont Health Monit*. 5960162. https://doi.org/10.1155/2024/5960162.
Scozzese, F., G. Leoni, and A. Dall'Asta. 2025. "HHT-based probabilistic model of prestressed bridges inferred from traffic loads." *Struct. Cont. Health Monit*. 2585257. https://doi.org/10.1155/stc/2585257.
Sekiya, H., K. Kimura, and C. Miki. 2016. "Technique for determining bridge displacement response using MEMS accelerometers." *Sensors* 16 (2): 257. https://doi.org/10.3390/s16020257.
Stone, H. S., B. Tao, and M. McGuire. 2003. "Analysis of image registration noise due to rotationally dependent aliasing." *J. Visual Commun. Image Represent.* 14 (2): 114–135. https://doi.org/10.1016/S1047-3203(03)00002-6.
Sun, L., Y. Li, and W. Zhang. 2020. "Experimental study on continuous bridge-deflection estimation through inclination and strain." *J. Bridge Eng.* 25 (5): 04020020. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001543.
Wang, J., and E. Olson. 2016. "Apriltag 2: Efficient and robust fiducial detection." In *Proc., Int. Conf. on Intelligent Robots and Systems*, 4193–4198. London, UK: IEEE.
Webb, G. T., P. J. Vardanega, and C. R. Middleton. 2015. "Categories of SHM deployments: Technologies and capabilities." *J. Bridge Eng.* 20 (11): 04014118. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000735.
Wenzel, H. 2009. *Health monitoring of bridges*. New York: Wiley.
Xing, L., W. Dai, and Y. Zhang. 2022. "Scheimpflug camera-based technique for multi-point displacement monitoring of bridges." *Sensors* 22 (11): 4093. https://doi.org/10.3390/s22114093.
Xu, Y., J. M. W. Brownjohn, and F. Huseynov. 2019. "Accurate deformation monitoring on bridge structures using a cost-effective sensing system combined with a camera and accelerometers: Case study." *J. Bridge Eng.* 24 (1): 05018014. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001330.
Yin, Y., Q. Yu, B. Hu, Y. Zhang, W. Chen, X. Liu, and X. Ding. 2023. "A vision monitoring system for multipoint deflection of large-span bridge based on camera networking." *Comput.-Aided Civ. Infrastruct. Eng.* 38: 1879–1891. https://doi.org/10.1111/mice.13050.
Yoneyama, S., and H. Ueda. 2012. "Bridge deflection measurement using digital image correlation with camera movement correction." *Mater. Trans.* 53 (2): 285–290. https://doi.org/10.2320/matertrans.I-M2011843.
Zhang, W., L. M. Sun, and S. W. Sun. 2017. "Bridge-deflection estimation through inclinometer data considering structural damages." *J. Bridge Eng.* 22 (2): 04016117. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000979.
Zhou, Y., J.-N. Hu, G.-W. Hao, Z.-R. Zhu, and J. Zhang. 2023. "Identification of influence lines for highway bridges using Bayesian parametric estimation based on computer vision measurements." *J. Bridge Eng.* 28 (12): 04023087. https://doi.org/10.1061/JBENF2.BEENG-6235.
Zona, A. 2021. "Vision-based vibration monitoring of structures and infrastructures: An overview of recent applications." *Infrastructures* 6 (1): 4. https://doi.org/10.3390/infrastructures6010004.