

Article

A Cost-Effective Template-Matching Vision System for Non-Contact Displacement Monitoring: Laboratory Validation Against Linear Variable Displacement Transducers

Valentina Picciano and Giuseppe Santarsiero * 

Department of Engineering, University of Basilicata, Via dell'Ateneo Lucano 10, 85100 Potenza, Italy; valentina.picciano@unibas.it

* Correspondence: giuseppe.santarsiero@unibas.it

Abstract

Structural health monitoring increasingly relies on non-contact optical techniques to overcome the limitations of contact sensors such as linear variable displacement transducers (LVDTs) and accelerometers. This paper presents the development and assessment of a vision-based displacement measurement platform, built in Python with OpenCV, implementing the Template Matching Method (TMM) for automatic tracking of user-defined regions of interest (ROIs) in video sequences. The platform was validated in a controlled laboratory environment through two configurations: a small circular target of 81 mm diameter and a much larger ROI on the flanges of an HEB 300 steel section, both instrumented with a reference LVDT. Video was acquired with a digital camera positioned frontally at 240 cm from the monitored elements, at 59.94 fps. Displacement time histories from template matching were compared against synchronous LVDT recordings using the coefficient of determination (R^2), mean absolute error (MAE), root mean square error (RMSE) and mean percentage residual. The small ROI yielded $R^2 = 0.9975$, MAE = 0.89 mm and RMSE = 1.06 mm, while the larger ROI yielded $R^2 = 0.9932$, MAE = 1.22 mm and RMSE = 1.51 mm. Both approaches required decontamination of the rigid-body motion of the supporting plate to which the LVDT was attached. Near-millimetre accuracy was obtained on both a marker-dominated and a texture-dominated ROI, at a fraction of the cost of alternative approaches reported in the literature. Beyond the present validation, the platform is offered as a self-contained, reusable research tool for displacement monitoring in other experimental configurations.

Keywords: structural health monitoring; computer vision; template matching; non-contact displacement measurement; cost-effective monitoring



Academic Editors: Miao Su, Zhiping Zeng, Min Wu and Fuming Wang

Received: 24 July 2026

Revised: 29 August 2026

Accepted: 30 August 2026

Published: 1 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Maintaining the structural integrity and functionality of infrastructure networks is one of the most technically and economically demanding challenges facing civil engineering. Bridges and viaducts occupy a strategic position within this landscape, as essential nodes for the mobility of people and goods whose vulnerability is intrinsically linked to material ageing, to traffic loads that increasingly exceed original design assumptions, and to the progressive action of environmental degradation. In reinforced concrete bridges, carbonation and chloride penetration trigger reinforcement corrosion and a loss of bond at the steel–concrete interface, with a macroscopic effect that consists of a progressive

reduction in global stiffness and safety margins. The intrinsically nonlinear evolution of these degradation phenomena is one of the main reasons why structural monitoring campaigns have become indispensable: detecting the advancement of damage in near-real time allows behavioural anomalies to be identified before they translate into irreversible states of distress [1].

Beyond damage detection, monitoring plays a decisive role in guiding the design of durability-enhancing and life-extending interventions on existing structures, which is often the only economically viable alternative to demolition and reconstruction for infrastructure of historical, architectural or strategic relevance. The retrofit of the half-joints of the Musmeci Bridge in Potenza through external prestressed tendons, for instance, illustrates how a targeted, monitoring-informed intervention can substantially prolong the service life of a structurally and architecturally significant bridge [2]. Within this broader perspective, Structural Health Monitoring (SHM) has progressively established itself not as an extraordinary verification tool but as a continuous knowledge-building process, aimed at shifting maintenance strategies from reactive to predictive paradigms grounded in the actual conservation state of the structure.

Among the physical quantities that can be observed, static and dynamic displacement plays a central role, since it directly and globally reflects structural stiffness, the cracking state and the possible onset of anomalous behaviour. Accurate in situ displacement measurement has historically been constrained by the operational limits of contact-based techniques. Linear variable displacement transducers (LVDTs) offer excellent metrological precision, but require scaffolding or mechanical supports anchored to stable external reference points, a condition that is often logistically and economically unsustainable in the presence of large spans, riverbeds, complex urban contexts or considerable heights. Dynamic monitoring based on accelerometers allows natural frequencies and mode shapes to be identified, but the reconstruction of displacement through double numerical integration of the acceleration signal is affected by drift and noise sensitivity that compromise its reliability over long durations [3,4].

It is precisely to overcome these technological barriers that research and engineering practice have increasingly turned toward non-contact (contactless) monitoring techniques. Among these, satellite radar interferometry has opened valuable perspectives for the surveillance of slow deformations at a territorial scale [5], while the advent of computer vision and digital image processing has opened revolutionary prospects at the structural and component scale, allowing continuous or discrete displacement fields to be extracted through optical sensors operating remotely from the structure. Two main families of vision-based approaches have emerged in the recent literature. The Template Matching Method (TMM) tracks a selected region of interest across a video sequence through similarity criteria, translating the geometric displacement of the tracked window in the image plane into a physical displacement once an appropriate scale factor is applied [4,6,7]. Phase-based motion magnification methods, by contrast, interpret motion as a variation in the local phase of the spatial components of the image signal, which makes them particularly suited to revealing micro-vibrations that would otherwise be buried in optical noise, at the cost of a more demanding computational and interpretative chain [8,9]. Both families belong to the broader class of contactless techniques and share the objective of converting a video sequence into quantitative information on structural motion, while responding to different observational logics [4].

The toolkit of computer-vision techniques applied to civil engineering extends well beyond displacement tracking, and increasingly includes deep learning-based approaches for the automated visual assessment of infrastructure components, such as the regional prioritisation of bridge guardrail conditions from open-source imagery [10] or crack de-

tection in civil infrastructure [11]. Non-contact, camera-based assessment is particularly valuable in post-earthquake emergency scenarios, where the rapid usability evaluation of a large number of buildings and structures must be carried out under severe time constraints and where the installation of dedicated contact sensors is rarely feasible; rapid visual survey procedures developed for post-earthquake usability assessment [12] and site-damage reconnaissance campaigns [13] both illustrate the operational demand for tools that can characterise structural condition quickly, remotely and without physical access to the structure, a demand to which vision-based displacement monitoring is directly responsive.

Vision-based displacement measurement has been applied with increasing success to in-service bridges, from medium-span post-tensioned concrete decks monitored under vehicular traffic [3] to steel or composite bridges observed during incremental launching construction [14], to static and dynamic monitoring campaigns aimed at correlating the measured response with the passage of moving loads [15]. Railway bridges, in particular, have motivated research toward sub-millimetre accuracy given the sensitivity of high-speed traffic to small dynamic displacements [16], while the increasing availability of urban surveillance and consumer-grade cameras has stimulated the development of self-calibration procedures able to correct perspective distortion and baseline drift in uncontrolled acquisition conditions [17]. Across this body of work, a recurring conclusion is that the accuracy of a vision-based measurement does not depend on the tracking algorithm alone, but on the entire acquisition and processing chain, including the camera-to-target distance, image resolution, frame rate, lighting stability and correctness of the pixel-to-metric calibration [4,18].

This paper focuses on the laboratory validation stage of a broader research activity aimed at developing and testing a vision-based displacement monitoring platform, built in Python and OpenCV, for both static and dynamic applications. The laboratory tests used for this validation consist of a quasi-static, displacement-controlled test on specimens of a Steel Plate Energy Absorption Device (SPEAD), a non-invasive, externally applied steel dissipative device developed for the seismic strengthening of existing RC frame structures [19,20]; the present tests were carried out within the broader experimental campaign aimed at characterising the cyclic response of the retrofitting element, which also relies on numerical models calibrated against the recorded displacement histories. While the complete research programme also included a field campaign on an in-service bridge, the present work is deliberately restricted to controlled laboratory tests, in which the displacement time histories reconstructed via the Template Matching Method are compared against a reference LVDT on two different targets: a small circular artificial marker and a broader region of a steel profile belonging to the test apparatus, HEB 300 section, directly on its natural surface texture even though including the circular target. This scope allows the metrological performance of the platform to be isolated from the additional uncertainties that are introduced by uncontrolled outdoor acquisition conditions, providing a clean benchmark of the accuracy that can be expected from the method under favourable, well-instrumented conditions, and a direct comparison between a small ROI in which the marker dominates the tracked frame and a much larger ROI in which the same marker occupies only a minor fraction of the frame, on the same test rig. The contribution of this paper is threefold: (i) a same-rig, same-sensor, same-marker quantitative comparison between a marker-dominated and a texture-dominated ROI, isolating the effect of target size and geometry from every other variable in the acquisition chain; (ii) a video-based rigid-body-motion correction procedure that isolates and removes the parasitic motion of the LVDT's own support directly from the recorded video, using the same camera and without any additional instrumentation; and (iii) a fully documented, reusable open-source processing platform, released together with its graphical user interface, that other researchers can apply to different structural elements and test setups without modification of its core

tracking logic. The remainder of this paper is organised as follows. Section 2 describes the processing platform, the working principle of the Template Matching Method, the laboratory specimen and experimental setup, and the statistical metrics used for validation. Section 3 presents the results obtained on the two target configurations. Section 4 discusses the findings in relation to the existing literature, and Section 5 draws the conclusions of this study.

2. Materials and Methods

2.1. Template Matching Method

The Template Matching Method is one of the most established approaches for non-contact displacement measurement based on computer vision. Its operating principle consists of selecting, in an initial frame, a small region of the image (the template) characterised by a sufficiently distinguishable intensity pattern, and of searching for the corresponding position of that pattern in the subsequent frames through similarity criteria. The variation in the template position in the image plane is interpreted as the in-plane motion of the observed portion of the structure, and is subsequently converted into physical units through a calibration procedure based on a known reference length within the field of view (Figure 1).

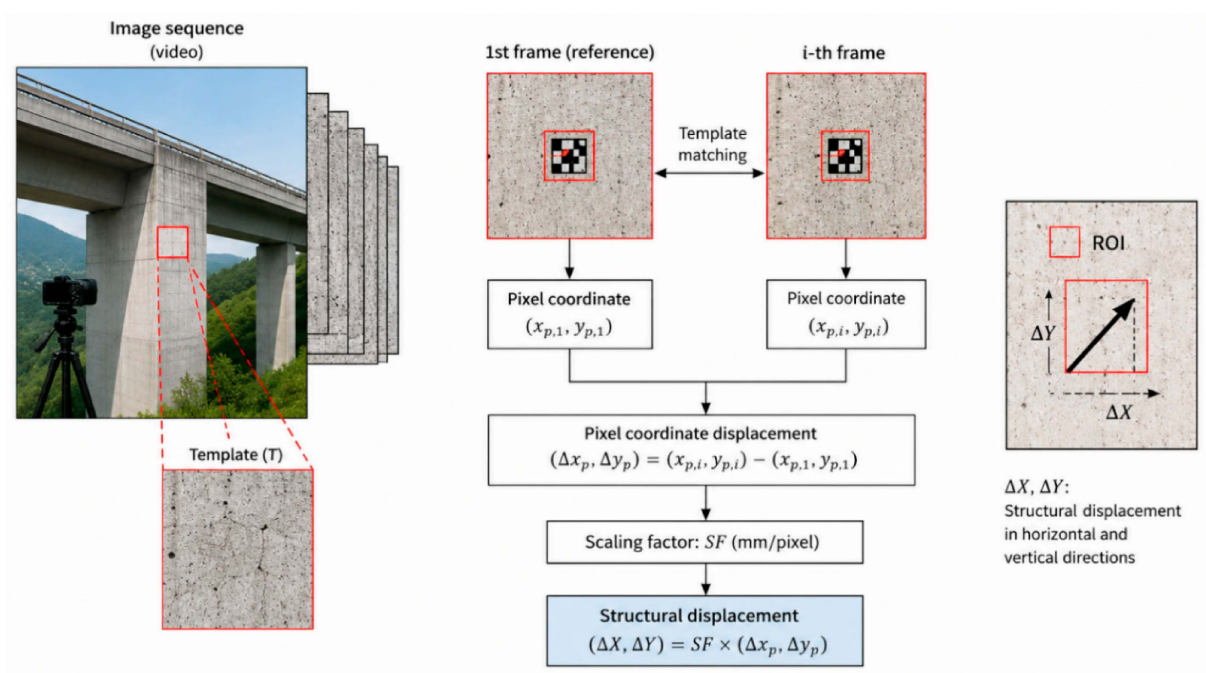


Figure 1. Schematic workflow of the vision-based, template-matching displacement measurement chain, from video acquisition to the physical displacement time history.

The reliability of the resulting measurement does not depend on the matching algorithm alone but on the entire acquisition and processing chain: the optical quality of the scene, the spatial resolution of the sensor, the frame rate, the choice and size of the target, the correctness of the calibration and the stability of the camera support. The method performs particularly well when the tracked region presents sharp, stable visual features and when the acquisition conditions are well controlled, but its accuracy can degrade when illumination changes, when contrast is lost, when motion blur is present, or when the camera itself is subject to unwanted movement. For this reason, the present study explicitly addresses two aspects of this metrological chain that are often left unexamined: the sensitivity of the method to target size and geometry, tested here through a small ROI in which the marker occupies most of the frame and a much larger ROI on the same structural

section in which the marker occupies only a minor fraction of it, and the correction of spurious rigid-body motion transmitted through the test setup, addressed through the decontamination procedure described in Section 2.3 [4,21,22]. A further, largely underexploited advantage of the vision-based approach is that the camera does not observe a single, isolated point of the specimen, as a contact transducer does, but the entire framed portion of the test apparatus at once. A single video recording therefore allows several regions of interest to be tracked simultaneously and retrospectively. For instance, the monitored structural element can be tracked along with the support on which the reference transducer is itself installed so that any parasitic or unwanted motion transmitted through the test rig can be isolated and subtracted from the structural response after the test, without having to anticipate and instrument every possible source of spurious displacement in advance.

2.2. Vision-Based Displacement Measurement Platform

The experimental application of vision-based monitoring was carried out through a dedicated processing platform developed in the Python (version 3.13.5) environment [23], with the objective of converting acquired video sequences into displacement time series that could be directly compared with traditional sensors. Python was chosen for its flexibility, reproducibility and compatibility with specialised libraries for numerical and image processing. The tracking core of the platform was implemented using the OpenCV library (version 4.13.0) [6], employed for frame management, definition of the regions of interest and application of the template matching algorithm. This architecture allows the tracking of the monitored point to be automated across the entire video sequence, reducing manual operator intervention and improving the consistency of the displacement extraction stage. The platform was deliberately kept independent of the specific specimen used in the present validation: it accepts any video file, any user-defined ROI and any reference length entered by the operator, so that it can be applied without modification to other displacement-monitoring tasks, and is offered here as a reusable research product in its own right, alongside the experimental findings. The cost-effectiveness of the platform rests only in part on the modest purchase price of its hardware (Table 1); more importantly, the camera, lens and tripod used here are general-purpose equipment already commonly available in structural and civil engineering practice, or readily reusable for other routine tasks such as visual inspection surveys, photogrammetric surveys and general site documentation, rather than being a piece of equipment purchased and dedicated solely to displacement monitoring. This distinguishes the proposed approach from dedicated instrumentation such as LVDTs, laser systems or robotic total stations, whose cost is not amortised over other uses. Table 1 summarises the approximate cost of the hardware components used in the present validation. The processing platform itself is built entirely on free, open-source components (Python, OpenCV), with no licence fee.

Table 1. Approximate cost of the hardware components used in the present validation.

Hardware Component	Approximate Cost (USD)	Reuse/Versatility
Canon EOS 90D camera body	~700	General-purpose DSLR; reusable for inspection surveys, photogrammetry, general documentation
Canon EF 17–40 mm f/4L USM lens	~450	General-purpose zoom lens; reusable across other photographic and inspection tasks
Tripod	~150	General-purpose support equipment; reusable for any camera-based task
Processing software (Python, OpenCV)	0 (open-source)	No licence fee; freely reusable and modifiable
Total	~1300	

A graphical user interface was developed to support the operator in each step of the analysis (Figure 2), organised into four panels covering video selection, template and scale definition, experimental data import, and analysis and comparison. The video panel lists the recordings available in the working directory and, once a file is selected, reads its frame rate and total frame count directly from the file metadata.

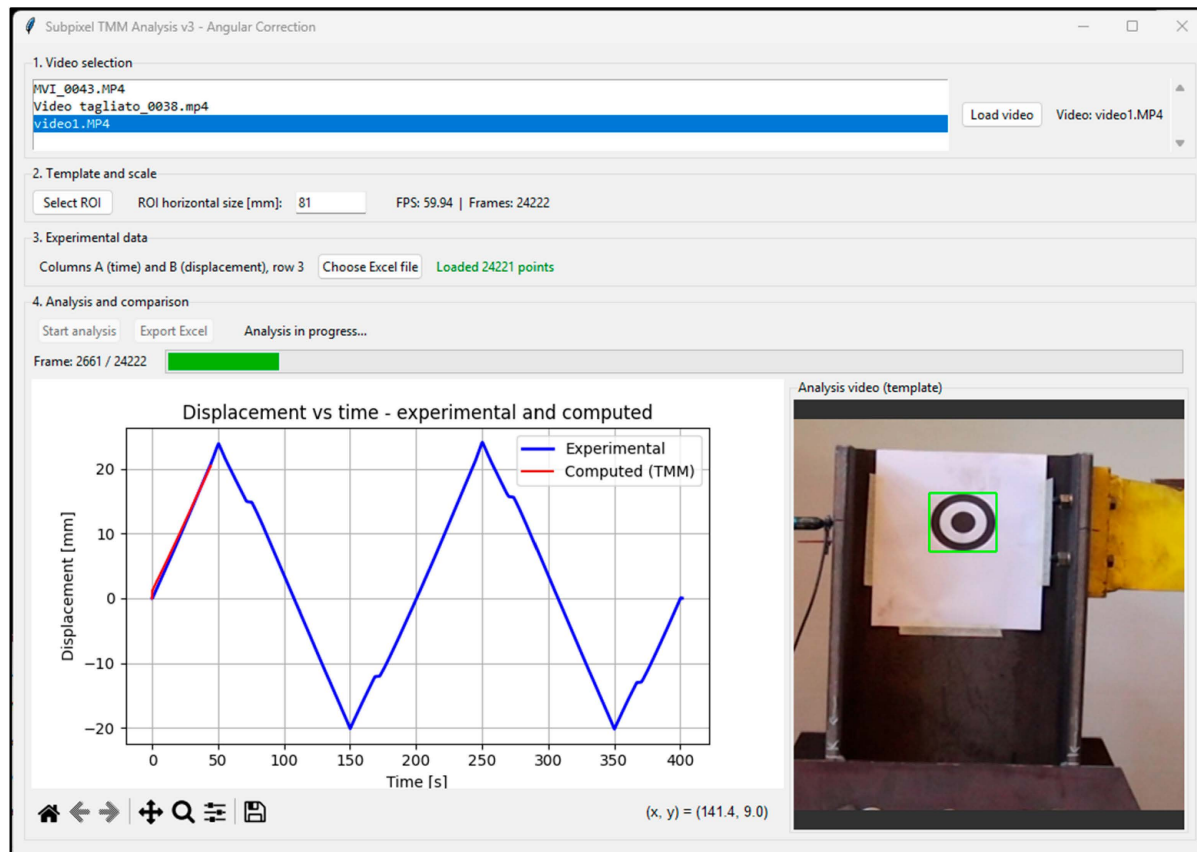


Figure 2. Graphical user interface of the vision-based processing platform, showing the video selection, ROI and scale definition, experimental data import, and the displacement-versus-time comparison panel.

Definition of the tracking template follows a two-stage selection procedure: the operator first draws a coarse rectangular region on the first video frame, which the software digitally magnifies by a factor of two, and then draws the actual template within this enlarged view, allowing the reference window to be positioned with greater precision than the original frame resolution would permit. The real-world horizontal size of the selected reference is then entered by the operator, from which the software computes the millimetre-per-pixel scale factor used to convert pixel displacements into physical units.

The analysis itself runs in a background thread, so that the interface remains responsive while each frame is processed. A progress bar and a frame counter report its advancement in real time. For every frame, the template is located within the magnified search window through normalised cross-correlation (OpenCV's `TM_CCOEFF_NORMED` criterion), and the position of the resulting correlation peak is further refined to sub-pixel accuracy by a two-dimensional parabolic interpolation of the correlation values around the peak, which is what allows the platform to resolve displacements markedly smaller than one pixel [24]. An embedded plot updates continuously as the analysis proceeds, overlaying the computed displacement–time curve on the imported experimental curve, while a synchronised preview panel displays the tracked template location on the current frame. No minimum correlation-value threshold is enforced for match acceptance: the global maximum of the

normalised cross-correlation surface is accepted as the match at every frame; this is considered acceptable for the short, controlled laboratory tests presented here, where lighting was stable and the target remained continuously visible throughout, but would need to be introduced for field applications where occlusion or lighting changes are more likely. The template is extracted once by the operator from the first frame and is never updated or re-acquired during the analysis, so tracking relies throughout on a single, fixed reference template; a static template avoids the risk of error accumulation ('drift') that a periodically updated template can introduce, at the cost of reduced robustness to gradual changes in target appearance over long recordings. The search window is not a small window re-centred frame-by-frame on the previous match, but the entire user-selected, digitally magnified region of interest, within which the correlation surface is computed at every frame. The sub-pixel offset of the correlation peak is computed through two-dimensional parabolic interpolation only where the local curvature of the correlation surface around the peak is non-negligible (the denominator of the parabolic fit exceeds 1×10^{-10} in both directions); where this condition is not met, the integer-pixel peak location is retained without sub-pixel correction. Once the analysis is complete, the paired experimental and computed time series, together with a comparison plot, can be exported to a spreadsheet file for further statistical analysis.

It is noted, for transparency, that a generative artificial intelligence tool was used solely to assist with English-language editing of this manuscript (grammar, phrasing and clarity); it played no role in the design, execution or analysis of the experiments described in this paper, nor in the interpretation of the results.

2.3. Laboratory Experimental Setup

The specimen used for the laboratory tests was the Steel Plate Energy Absorption Device (SPEAD), an external, non-invasive steel dissipation system developed to enhance the seismic capacity of existing RC joints while minimising the impact of the intervention on the building and its occupants [19]. The device dissipates seismic energy through the hysteretic behaviour of steel plates applied externally to the joint panel, and its installation does not require the removal of finishes or the interruption of the building's use. The device was tested under quasi-static loading applied through a hydraulic actuator, in order to characterise its force–displacement response and to provide displacement time histories that could subsequently support the calibration of numerical models of the device, which will allow the numerical modelling of SPEAD-retrofitted joints [20]. It is within this experimental context that the vision-based platform described in Section 2.2 was deployed alongside the reference LVDT measurements.

The experimental tests were carried out in the structural testing hall of the Laboratory of Structures of the University of Basilicata, where the portal frame test rig was instrumented with a reference LVDT (± 100 mm stroke) whose displacements were acquired using a 50 Hz sampling strategy. Specifically, a linear potentiometric displacement transducer (Leane, model SLS190, 100 mm electrical stroke, guaranteed linearity $\pm 0.15\%$ F.S., hysteresis/repeatability < 0.01 mm) was used as the reference sensor. Displacements were simultaneously recorded with a Canon EOS 90D digital camera, equipped with a Canon EF 17–40 mm f/4L USM zoom lens, positioned in front of the specimen (Figure 3). The camera operated in Program AE (automatic exposure) mode, with the lens aperture fixed at f/4 and ISO set to Auto (base ISO 100); video was encoded in H.264/AVC (Canon 'Light (IPB)' compression profile) at an average bitrate of 60.2 Mbps. The camera was mounted on a tripod and kept stationary for the entire duration of each test, at a height of 131 cm above the floor (Figure 4) and at a distance of 240 cm from the monitored target (Figure 5); the target itself was positioned orthogonally to the optical axis of the camera, at the same

height of 131 cm, in order to maximise the orthogonality between the optical axis and the plane of motion. This frontal configuration was adopted deliberately, since out-of-plane inclination of the camera with respect to the target plane can alter the correspondence between real and pixel-measured displacement, introducing parallax errors that a purely frontal setup minimises. The camera operated at a frame rate of 59.94 fps with the lens set to a focal length of 20 mm, acquiring full HD (1920×1080 pixel) video sequences. These parameters were selected to guarantee a temporally dense sequence, adequate for the subsequent reconstruction of the displacement time history. One single video recording was performed to capture the entire experimental test, throughout its full duration.



Figure 3. Laboratory test setup: digital camera positioned frontally with respect to the instrumented portal-frame specimen, at a fixed distance and height from the monitored target.



Figure 4. Verification of the camera height (131 cm) above the floor, matching the height of the monitored target, measured with a graduated staff placed in front of the lens.

Two different reference targets were used in order to assess the robustness of the Template Matching Method with respect to target size, geometry and surface characteristics, and, more specifically, to compare a small ROI, in which the marker dominates the tracked frame, against a much larger ROI in which the same marker occupies only a minor fraction of the frame. This

kind of validation serves to recognise the effectiveness of the template matching approach in situations in which a specific target cannot be applied (e.g., due to accessibility constraints) and the natural texture of the structure is the main target to be used.



Figure 5. Verification of the camera-to-target longitudinal distance (240 cm), measured with a tape stretched in front of the lens.

The first reference consisted of a circular target of 81 mm diameter, printed and applied directly to the tested structure (Figure 6): a dedicated, high-contrast artificial marker of this kind constitutes a simple, well-delimited and easily recognisable geometric reference, but its reduced size makes the measurement more sensitive to localisation error, since even a small inaccuracy in the identification of the centroid or the edge of the target can produce a non-negligible percentage deviation, particularly when the displacements to be measured are of the order of a few millimetres. Under this configuration, the quality of the result depends significantly on image sharpness, camera stability and the correct identification of the geometric reference used for the pixel-to-millimetre conversion. The 81 mm diameter was chosen to keep the marker small relative to the 300 mm width of the HEB 300 flanges' distance, so that, when the same marker was framed within the much larger ROI used in the second configuration, it would occupy only a limited portion of the tracked area rather than dominating it.

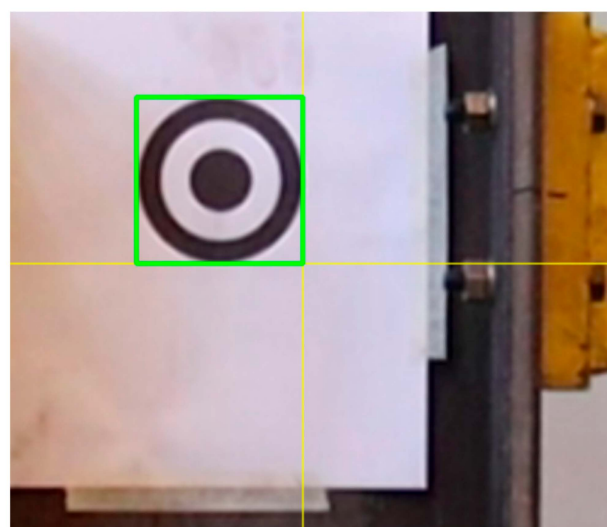


Figure 6. Region of interest (ROI) selection for the circular target (diameter 81 mm).

The second reference consisted of a considerably larger ROI including the flanges of an HEB 300 steel section, with a reference length of 300 mm along the tracked direction (Figure 7). This ROI still included the same circular marker used in the first configuration, since the marker itself remained affixed to the flange; it was, however, deliberately sized so that the 81 mm marker occupied only a minor fraction of the much larger tracked area, with the surrounding natural steel surface providing most of the visual content used for matching. The larger surface extent of this ROI offers the algorithm a wider region containing more usable visual information for recognising the template position in subsequent frames, which tends to reduce sensitivity to localised pixel errors and to make the matching less dependent on a single, small geometric feature. The steel surface surrounding the marker, however, cannot be treated as an ideal reference: reflections, contrast variations, imperfectly defined edges and micro-variations in surface texture can influence the correlation calculation, so that the definition of the ROI on the HEB 300 profile had to avoid visually ambiguous or unstable areas. The comparison between the two configurations therefore provides a direct, same-rig assessment of the accuracy gap between a small ROI in which the marker dominates the tracked area and a much larger ROI in which the same marker contributes only a minor share of the visual information, which is of practical relevance whenever a marker cannot be excluded from the field of view but its influence on tracking needs to be limited.

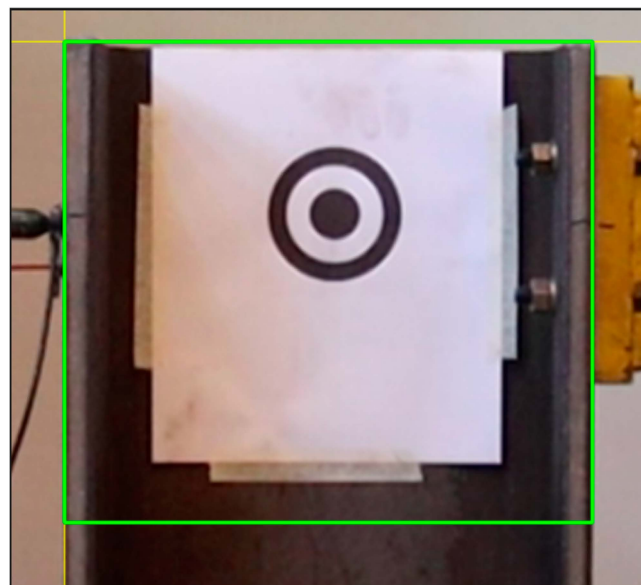


Figure 7. Region of interest (ROI) selection on the flange of the HEB 300 steel section (reference length 300 mm).

A specific methodological step was required to compare LVDT measurements with the vision-based ones. The vision-based method evaluates the absolute displacement of the tracked region with respect to the fixed camera reference frame, whereas the LVDT evaluates a displacement that is relative to its own point of installation, corresponding to the 42.5 mm thick plate visible in Figure 8. This plate is not rigidly connected to the laboratory floor but is itself part of the test rig, so it is not a perfectly stationary reference: any small displacement of the plate is implicitly embedded in the LVDT reading without being distinguishable from the actual response of the beam. To make the two measurements comparable on a consistent basis, the Template Matching Method was therefore applied not only to the ROIs but also, independently, to the 42.5 mm plate itself, within the very same video recording (Figure 8) and without any additional instrumentation. The displacement time history obtained for the plate was then subtracted from that obtained for the beam, so

as to purify the absolute displacement of the ROI from the rigid-body motion of the plate and to obtain a quantity that is directly comparable with the relative displacement measured by the LVDT. The displacement of the tracked target and the displacement of the supporting plate were computed independently, through two separate template-matching elaborations run on the same video recording, each with its own template and its own tracked position over time; only afterwards was the plate displacement time history subtracted from the target displacement time history, according to $\delta_{\text{corr}}(t) = \delta_{\text{ROI}}(t) - \delta_{\text{plate}}(t)$, to obtain the corrected result compared with the LVDT. Both elaborations share the same camera, frame and calibration reference, so no separate spatial calibration was required between the two tracked regions; the two targets lie at comparable depth within the frontal field of view, which is noted here as a simplifying assumption. This decontamination step is essential for a correct comparison between the video-based and the experimental datasets, since the former describes an absolute displacement while the latter is inherently a relative one, referred to a support that is not perfectly fixed. More generally, this capability illustrates a distinctive strength of vision-based monitoring with respect to point sensors: a single camera view can simultaneously capture the response of the monitored element and of the surrounding parts of the test setup that may introduce unwanted or parasitic deformation, allowing such contributions to be identified and removed in post-processing rather than having to be prevented, in advance, through a stiffer or more elaborate physical setup.

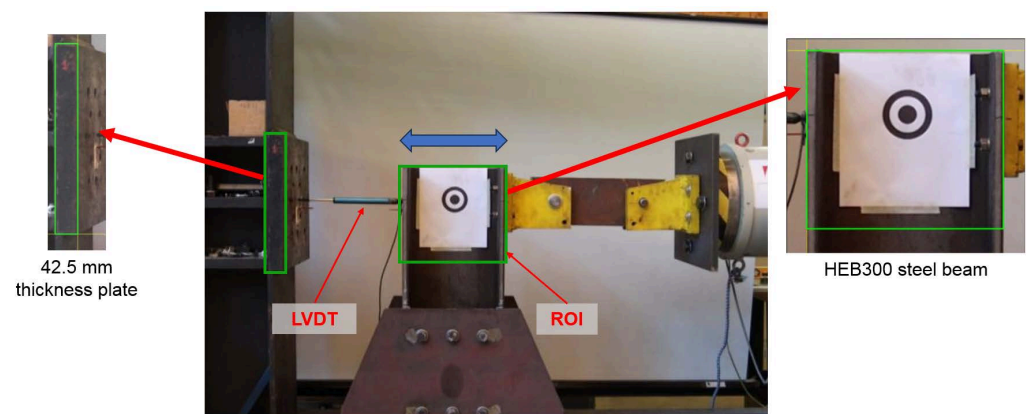


Figure 8. Overview of the HEB 300 test configuration, showing the LVDT installation, the region of interest (ROI) tracked on the HEB 300 flange, and the 42.5 mm thick supporting plate independently tracked with the Template Matching Method for the decontamination of its rigid-body displacement.

2.4. Evaluation Metrics

Prior to the comparison of the two monitoring approaches, the TMM signal was processed to make it synchronous with respect to the LVDT one. In fact, the two time series were recorded with different acquisition frequencies, namely 50 Hz and 59.94 Hz for the LVDT and the video sequence, respectively. The two series were aligned on a common, physically defined time base by computing the absolute time of each sample from its own acquisition frequency (50 Hz for the LVDT, 59.94 Hz for the video-derived TMM series), rather than by resampling one series onto the other's discrete grid beforehand. The starting instant (t_0) of each series was set at the first non-zero displacement recorded in that series, and this instant was used as the common time-alignment reference between the LVDT and TMM records; no additional time-shift correction beyond this t_0 -alignment was applied. For the quantitative comparison, the LVDT record was subsequently linearly interpolated onto the TMM sampling instants, so that a directly comparable pair of values was available at every video frame.

The quantitative comparison between the LVDT reference measurement and the displacement time history obtained through template matching was based on four comple-

mentary indicators: the mean absolute error (MAE), the root mean square error (RMSE), the coefficient of determination (R^2) and the mean percentage residual. Denoting the LVDT reading at the i -th sample by y_i , the corresponding vision-based estimate by $\hat{y}_i$, the sample mean of the reference series by $\bar{y}$ and the number of observations by n , the metrics are defined as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

$$R^2 = 1 - \left[\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \right] \quad (3)$$

The MAE provides an immediate measure of the average discrepancy between the two series, expressed in the same physical unit as the displacement (millimetres), without regard to the sign of the error. The RMSE assigns a larger weight to the largest deviations, since residuals are squared before averaging; consequently, comparing MAE and RMSE indicates whether the error is fairly uniformly distributed or whether isolated, more pronounced deviations are present. When RMSE is only moderately larger than MAE, the error tends to be regular and free of anomalous spikes. The coefficient of determination R^2 measures the extent to which the vision-based series reproduces the trend of the reference series, with values close to unity indicating an excellent agreement between the two datasets. In addition, four further indicators were computed to more fully characterise the agreement between the two series, following a Bland–Altman approach: the mean bias (MBE), the standard deviation of the residuals (SD), the maximum absolute error, and the mean percentage error. The mean bias and its standard deviation are defined as $MBE = (1/n) \sum (\hat{y}_i - y_i)$ and $SD = \sqrt{[(1/(n-1)) \sum (e_i - MBE)^2]}$, where $e_i = \hat{y}_i - y_i$ is the residual of the i -th sample; a positive MBE indicates that the vision-based estimate systematically exceeds the LVDT reading. The 95% limits of agreement, $MBE \pm 1.96 \cdot SD$, and the maximum absolute error, $\max_i |e_i|$, complete the characterisation of the residual distribution beyond the correlation-based R^2 . The mean percentage error is defined as $e_i(\%) = [(\hat{y}_i - y_i)/y_i] \times 100$, averaged over all samples for which $|y_i| \geq 1$ mm; samples with $|y_i| < 1$ mm, in the neighbourhood of the zero crossing of the LVDT signal, were excluded from this indicator only, since the denominator tends to zero while the numerator does not vanish at the same instants, which would otherwise produce artificially large or undefined percentage values that are not representative of the actual tracking error. These samples remain fully included in every other indicator (MAE, RMSE, MBE, SD, maximum absolute error), which do not share this singularity.

3. Results

3.1. Circular Target

For the circular target configuration, the comparison between the displacements obtained through template matching and those recorded by the LVDT (Figure 9) produced a coefficient of determination $R^2 = 0.9975$, with $MAE = 0.8905$ mm and $RMSE = 1.0570$ mm. These results indicate an extremely high correlation between the two measurement series (Figure 10) and confirm the ability of the algorithm to correctly reconstruct the trend of the displacement. The limited value of the MAE shows that the average absolute error remains below one millimetre, while the RMSE, only slightly higher, indicates that no particularly marked isolated deviations are present. The mean bias was $MBE = 0.660$ mm, with a standard deviation of the residuals $SD = 0.826$ mm (95% limits of agreement: -0.958 mm to 2.278 mm) and a maximum absolute error of 2.719 mm, confirming that some anomalous deviation is present only at highest values of measurement. The mean percentage error,

computed only for samples with $|\text{LVDT}| \geq 1$ mm to avoid the zero-crossing singularity described in Section 2.4, was -3.24% ; the difference in sign between this relative indicator and the small positive MBE reported above reflects the different weighting of small- and large-displacement samples between an absolute and a relative error metric, rather than a contradiction in the underlying residuals [21].

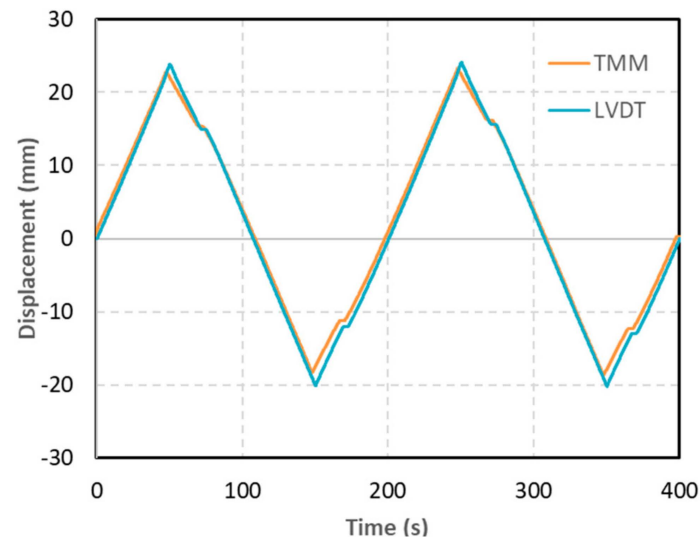


Figure 9. Time-history overlap between the LVDT reference measurement and the Template Matching Method (TMM) estimate for the circular target.

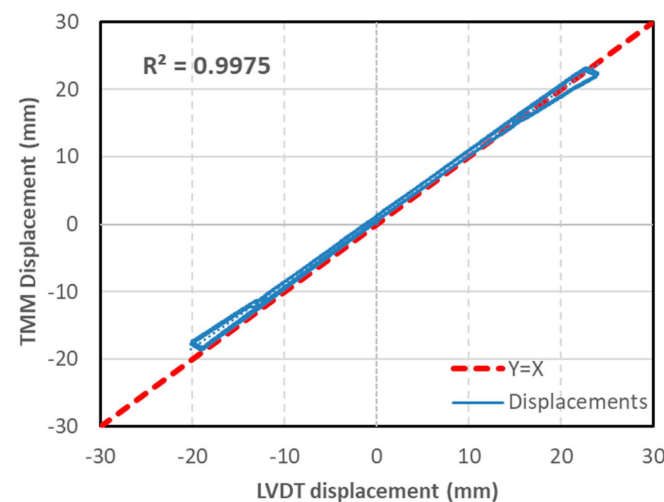


Figure 10. Vision-based (TMM) versus LVDT-measured displacement for the circular target ($R^2 = 0.9975$).

Figure 11 presents the corresponding Bland–Altman plot for the circular-target configuration. Across most of the tested range, away from the extremes of travel, the residuals remain comfortably within the 95% limits of agreement (-0.958 mm to 2.278 mm) and trace two closely spaced branches, consistent with the repeated loading/unloading cycles, with a mild decrease in the difference as the mean displacement increases from negative to positive. Close to the two extremes of the displacement range, however, the residual briefly exceeds these limits: near the most negative mean displacement (≈ -19 mm), it peaks at about 2.7 mm, above the upper limit of agreement, and near the most positive mean displacement ($\approx +23$ mm), it drops sharply to about -2 mm, below the lower limit. These brief excursions coincide with the reversal points of the cyclic test, where the direction of motion changes, and indicate that the largest tracking discrepancies are concentrated

at the extremes of travel rather than being uniformly distributed across the range. A few additional brief, sharp deviations are visible near zero mean displacement, coinciding with the zero-crossing instants excluded from the percentage-error indicator (Section 2.4).

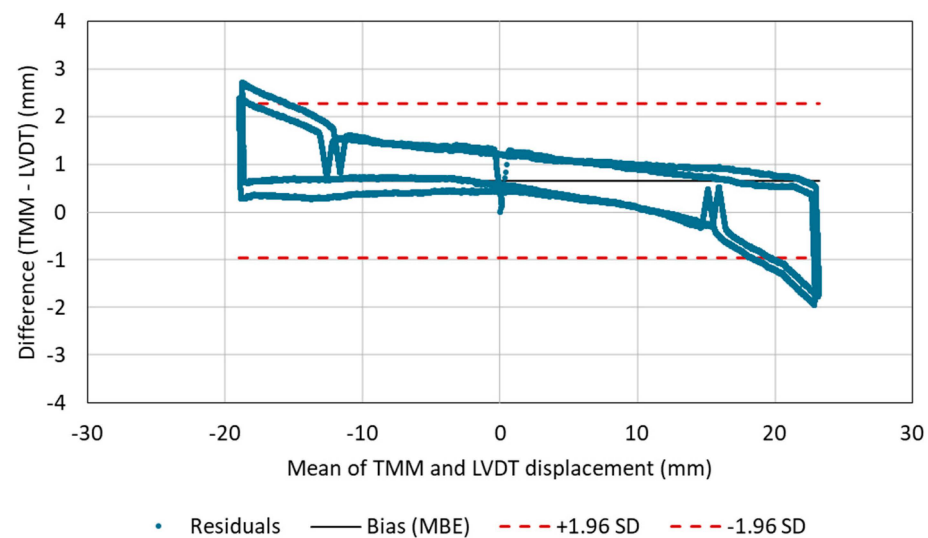


Figure 11. Bland–Altman plot, circular target (mean vs. difference, with bias and 95% limits of agreement lines).

The result obtained on the circular target shows that template matching can achieve good accuracy even on geometric references of limited size, provided that the target is well contrasted, correctly framed and stable throughout the acquisition. At the same time, the reduced size of the target is a critical aspect from a metrological point of view: a small localisation error in the centroid or the edge of the marker can produce a percentage deviation that is proportionally more significant than it would be on a larger target. This is consistent with the literature on vision-based monitoring, in which the accuracy of the measurement depends significantly on spatial resolution, on the millimetre-per-pixel ratio, on calibration quality and on ROI stability, and in which the camera must be regarded not simply as an acquisition device but as an actual optical sensor whose precision depends on the acquisition geometry and on the ability of the algorithm to keep the tracking of the target stable over time [4].

3.2. HEB 300 Steel Section

For the HEB 300 configuration, the comparison (Figure 12) again produced very good results, although with a slightly higher average error. The coefficient of determination was $R^2 = 0.9932$, the MAE was 1.2214 mm and the RMSE was 1.5133 mm. These figures confirm that the vision-based method retains a high capacity to follow the experimental trend even on a wider and structurally more complex geometry (Figure 13). The mean bias was $MBE = -0.868$ mm, with a standard deviation of the residuals $SD = 1.240$ mm (95% limits of agreement: -3.298 mm to 1.562 mm) and a maximum absolute error of 3.411 mm. The mean percentage error, computed only for samples with $|LVDT| \geq 1$ mm, was -5.59% , consistent in sign with the negative MBE reported above.

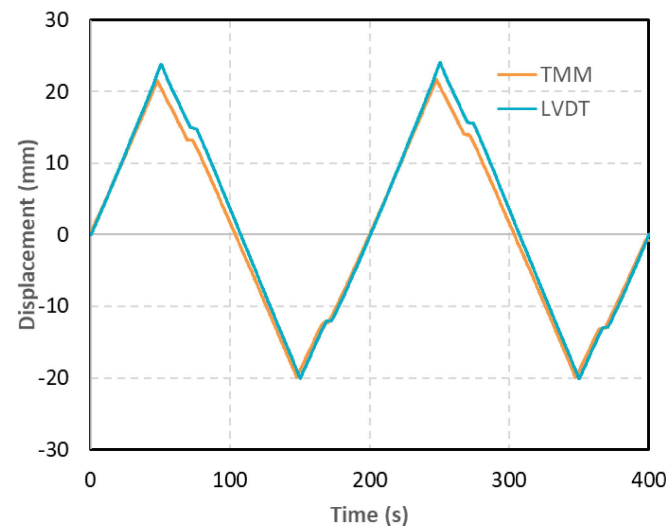


Figure 12. Time-history overlap between the LVDT reference measurement and the Template Matching Method (TMM) estimate for the HEB 300 section.

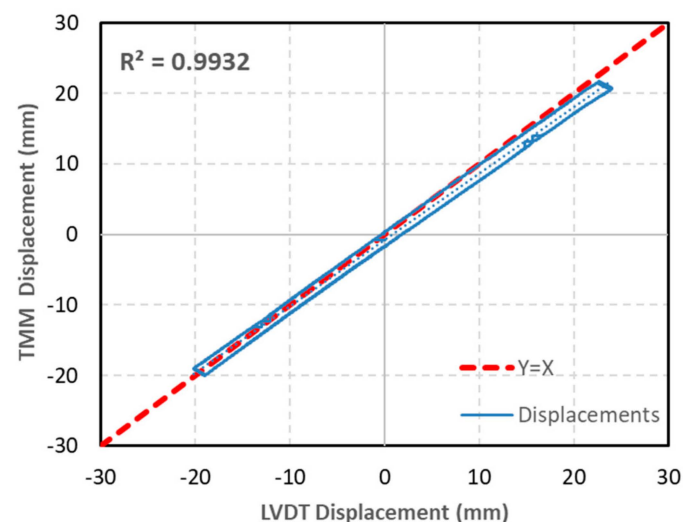


Figure 13. Vision-based (TMM) versus LVDT-measured displacement for the HEB 300 section ($R^2 = 0.9932$).

Figure 14 shows the same type of plot for the HEB 300 configuration. The pattern looks like the circular target: two branches from loading and unloading, but with more space between them here, matching the larger standard deviation in Table 2. The difference goes down as the mean displacement goes up. Here, almost all points stay inside the 95% limits (-3.298 mm to 1.562 mm), with only one small exception: at the far positive end ($\approx +22$ mm), the residual drops to about -3.4 mm, just below the lower limit. This is a smaller and more contained deviation than the ones seen at both ends of the range for the circular target.

The moderate increase in MAE and RMSE with respect to the circular target should not be interpreted as a substantial deterioration of the procedure, but rather as a consequence of the greater complexity of the measurement chain in this configuration. It is worth stressing that, although the same circular marker remained within this wider ROI, it occupied only a minor fraction of the tracked area relative to the surrounding steel surface; the resulting R^2 of 0.9932 is therefore largely a measure of the accuracy achievable when tracking is dominated by the natural surface texture rather than by an idealised, high-contrast target, even though the marker was not excluded from the frame. Overall, both configurations

returned R^2 values above 0.99 and average errors of the order of one millimetre, confirming that the accuracy of the platform does not depend exclusively on the presence or the size of a dedicated target, but on the interaction between image scale, contrast, texture quality, camera stability, the nature of the measured displacement and the correctness of the calibration procedure.

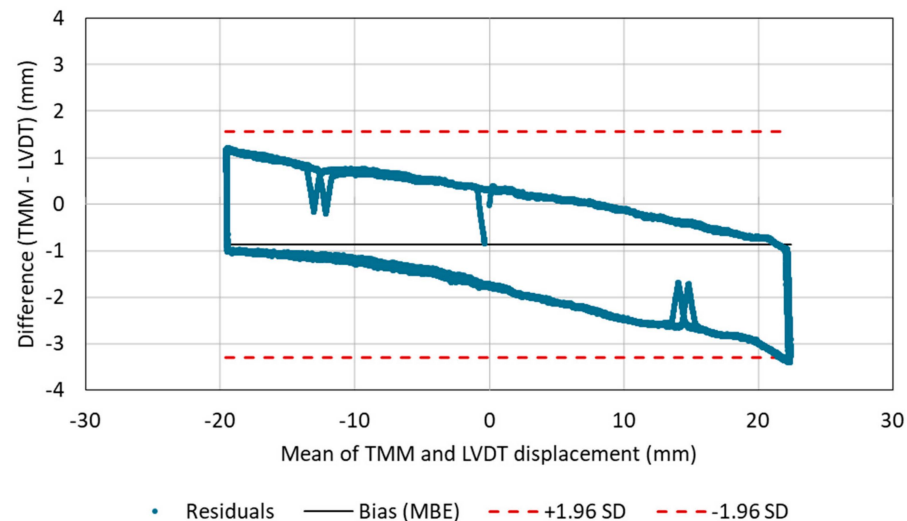


Figure 14. Bland–Altman plot, HEB 300 section (mean vs. difference, with bias and 95% limits of agreement lines).

Table 2. Comparison of accuracy metrics obtained with the small, marker-dominated ROI (circular target) and the much larger, texture-dominated ROI on the HEB 300 flange.

Metric	Circular Target	HEB 300 Steel Section
Reference length	81 mm (printed circular marker)	300 mm
Coefficient of determination, R^2	0.9975	0.9932
Mean absolute error, MAE (mm)	0.8905	1.2214
Root mean square error, RMSE (mm)	1.0570	1.5133
Mean percentage error ($ LVDT \geq 1$ mm)	−3.24%	−5.59%
Mean bias, MBE (mm)	0.660	−0.868
Standard deviation of residuals, SD (mm)	0.826	1.240
Maximum absolute error (mm)	2.719	3.411

3.3. Comparison of the Two Configurations

Table 2 summarises the accuracy metrics obtained for the two target configurations, allowing the small, marker-dominated ROI and the much larger, texture-dominated ROI to be compared directly on a common metrological basis.

The comparison, considering the mean bias, the standard deviation of the residuals and the maximum absolute error requested above, confirms that the small, marker-dominated ROI (circular target) achieved superior accuracy on every indicator in Table 2: a higher R^2 , a lower MAE, a lower RMSE, a smaller magnitude of the mean bias, a smaller standard deviation of the residuals and a smaller maximum absolute error than the much larger ROI on the HEB 300 flange. On strictly metrological grounds, the small ROI should therefore be regarded as the more accurate of the two approaches, consistent with the higher contrast and better-defined geometry of a marker that occupies most of the tracked frame. A plausible explanation for the consistently larger bias and dispersion observed on the HEB

300 configuration is that its much larger, texture-dominated ROI is more exposed, over its wider area, to lighting gradients, minor out-of-plane effects and local surface imperfections than the smaller, high-contrast circular marker, whose tracked area is small enough to remain nearly uniform under the same illumination. Moreover, the HEB300 ROI is prone to rotating during the application of the controlled displacement, introducing a distortion of the ROI itself, which could affect the template matching process. This advantage, however, needs to be weighed against a different practical constraint: keeping the marker large relative to the tracked area requires a tightly cropped ROI, which offers less resilience to local disturbances than a wider frame. The much larger ROI on the HEB 300 section, in which the same marker occupied only a minor fraction of the tracked area, reproduced the LVDT measurement with an accuracy only slightly lower than that of the small ROI, which indicates that relying predominantly on the surrounding natural surface, rather than on a marker that dominates the frame, is a viable alternative whenever a large, well-contrasted marker cannot be applied. The choice between the two approaches is therefore not a matter of one being unconditionally better than the other, but of matching the ROI size to the constraints of the specific monitoring task.

4. Discussion

The laboratory results obtained on the two target configurations allow some general considerations to be drawn on the metrological behaviour of template-matching-based displacement measurement. The comparison between the small, marker-dominated ROI and the much larger ROI on the HEB 300 section highlights a trade-off that is well known in the vision-based monitoring literature but is rarely quantified on the same experimental rig, with the same reference sensor, and, as in this case, using the very same physical marker framed at two different scales: a tightly cropped ROI, in which the marker occupies most of the tracked area, is more sensitive to pixel-level localisation error, because any inaccuracy in the definition of its centroid or edge has a proportionally larger effect on the measured displacement, whereas a much larger ROI, in which the same marker contributes only a minor share of the visual content, draws more on the surrounding natural texture and is therefore more resilient to local disturbances. Both configurations nevertheless returned R^2 values in excess of 0.99, indicating that the platform preserves a high predictive capability regardless of whether the tracked frame is dominated by the marker or by the surrounding natural surface, as a result of direct practical relevance, since existing structural components in the field very rarely offer the operator the possibility of applying a marker large enough to dominate a tightly cropped, high-precision ROI.

A further methodological aspect concerns the conceptual difference between the two measurement modalities being compared. The vision-based system returns an absolute displacement of the observed portion of the structure, obtained from the displacement of the template in the image plane and subsequently converted into physical units, whereas the LVDT measures a relative displacement, tied to its own installation configuration and mechanical reference system. This difference explains why residuals between the two series can appear even in the presence of a very high overall correlation, and it is not necessarily an indication of an error in the video-based method but rather a structural difference between two distinct measurement principles. The decontamination of the plate displacement was specifically introduced to reduce the discrepancy by isolating the deformation contribution that can actually be attributed to the monitored target, and it was made possible only because the camera, unlike the LVDT, observed the specimen and the supporting plate simultaneously within the same field of view. This is arguably one of the most practically relevant outcomes of the present study: a vision-based system does not merely replace a point sensor with another point sensor placed at a distance, but effectively

transforms the acquisition into a multi-point measurement of the entire test apparatus, preventing any unwanted or parasitic deformation component.

The findings are consistent with the broader literature on vision-based structural displacement measurement. Feng et al. [4] and Dong and Catbas [7] have both emphasised that the accuracy of contactless displacement sensing is governed by the entire acquisition chain rather than by the tracking algorithm in isolation, a conclusion directly supported by the sensitivity observed here between target size and measurement error. Lee et al. [18] highlighted the importance of an adaptive and robust ROI in field applications, particularly when the target risks becoming less distinguishable under unfavourable lighting conditions; while the present study was carried out under controlled laboratory illumination, the marked sensitivity of the small, marker-dominated ROI to localisation accuracy, together with the good performance obtained with the much larger, texture-dominated ROI on the HEB 300 section, suggests that such adaptive strategies, capable of exploiting a wider field of view when a large, high-contrast marker cannot be applied, would become even more relevant when the method is transferred to outdoor, in-service conditions. More broadly, the agreement achieved here on both a small, marker-dominated ROI and a much larger, predominantly natural steel surface supports the growing body of evidence that template matching, despite its conceptual simplicity, can deliver a metrologically credible alternative or complement to conventional contact instrumentation when the acquisition geometry is properly controlled [3,14,15]. The same non-contact, remote-observation principle that proved effective here in a controlled laboratory setting is also what makes vision-based techniques attractive for the rapid, installation-free structural assessment that is needed in post-earthquake emergency scenarios, where time and site accessibility are severely constrained [12,13].

Placing these figures against the wider body of recent literature helps to establish whether the accuracy achieved here can be regarded as satisfactory, and at what cost it was obtained. Marendić et al. [25] validated a UAV-based homography method against LVDT and robotic total station measurements on a full-scale bridge under traffic loading and reported RMSE values below 1.4 mm, with a stationary tripod-mounted camera achieving an RMSE below 0.5 mm under otherwise similar imaging conditions; in a related study on displacement measurement during structural load testing, the same research group combined UAV and tripod-mounted platforms, a robotic total station and multi-image bundle block adjustment to report an RMSE of 1.21 mm and a mean difference of 0.34 mm between photogrammetric and LVDT displacements [26]. Wu et al. [27], tracking natural bridge features with a moving-window template matching algorithm supported by a deep learning segmentation model to isolate the bridge region, reported a relative error with respect to LVDT measurements of 1–3% on a laboratory beam model. The MAE and RMSE obtained in the present study (0.89–1.22 mm and 1.06–1.51 mm, respectively) and the associated mean percentage residuals (in the range 3.24–5.59%) fall within, or close to, these ranges, yet were achieved with a markedly simpler and lower-cost measurement chain: a single stationary digital camera and an open-source Python/OpenCV processing pipeline, with no UAV platform, no robotic total station, no multi-camera bundle adjustment and no additional laser-assisted calibration (see comparison in Table 3). In this sense, the platform developed here at a total cost of about 1300\$, can be regarded as a cost-effective alternative that delivers accuracy competitive with substantially more instrumented, more expensive vision-based systems, which is precisely the combination of characteristics that makes it attractive for routine laboratory testing and for preliminary field deployments where the budget or the logistics required by UAV or total-station assisted approaches are not justified. A further reduction in the residual error, if required, would likely call for exactly this kind of additional hardware or more elaborate processing, at a corresponding increase

in cost and operational complexity. It is worth stressing that this cost advantage is not simply a matter of a lower purchase price: unlike a dedicated LVDT, laser system or robotic total station, which are single-purpose instruments whose cost is entirely attributable to the displacement-monitoring task, the camera, lens and tripod used here are general-purpose equipment that a structural testing laboratory is likely to already own or to use for other purposes, such as documentation, inspection or photogrammetric surveys. The relevant cost comparison is therefore not simply hardware price against hardware price, but the marginal cost of adding a displacement-monitoring capability to equipment that already serves other purposes, against the cost of acquiring and maintaining dedicated, single-purpose instrumentation.

Table 3. Comparison of the present study with related vision-based displacement-monitoring studies discussed in Section 4.

Study	Methodology	Target/Features Tracked	Reference Sensor	Reported Accuracy	Equipment/Cost Profile
Present study	Single stationary camera, Template Matching Method (TMM), OpenCV	Small circular marker; large texture-dominated ROI (same rig, both tested)	LVDT	RMSE 1.06–1.51 mm; MAE 0.89–1.22 mm; R^2 0.993–0.998	Single digital camera + tripod, open-source software; no UAV, no total station
Marendić et al. [25]	UAV-based and tripod-mounted homography	Structural targets, full-scale bridge under traffic	LVDT and robotic total station	RMSE < 1.4 mm (UAV); RMSE < 0.5 mm (tripod)	UAV platform + robotic total station
Marendić et al. [26]	UAV and tripod photogrammetry, multi-image bundle adjustment	Structural targets, load-testing campaign	LVDT and robotic total station	RMSE 1.21 mm; mean difference 0.34 mm	UAV + robotic total station + multi-camera bundle adjustment
Wu et al. [27]	Moving-window template matching with deep learning segmentation	Natural bridge features (no artificial marker)	LVDT	Relative error 1–3% (laboratory beam model)	Single camera + deep-learning segmentation model

The present study is limited to a controlled laboratory environment, with a fixed, short camera-to-target distance, stable illumination and a stationary camera support; these are favourable conditions that are not always reproducible during in-service monitoring of real bridges, where longer acquisition distances, environmental disturbances and the absence of dedicated, high-contrast targets can be expected to increase the measurement uncertainty relative to what has been reported here. The laboratory results should therefore be regarded as an upper bound on the achievable accuracy of the platform, and as a necessary metrological baseline against which any subsequent field application of the same processing chain can be benchmarked. Beyond the laboratory validation reported here, the same platform lends itself to several structural-health-monitoring scenarios that do not require a permanently installed contact sensor. In periodic, discontinuous monitoring campaigns on existing bridges, a camera can be temporarily set up to track one or more ROIs during a scheduled inspection visit, without the scaffolding or anchored supports that a contact transducer would require. In post-earthquake usability screening, the same rapid, installation-free deployment allows displacement or crack-opening measurements to be obtained on multiple structures within the severe time constraints typical of an emergency response. Finally, because a single video recording can be used to track several regions of interest simultaneously and retrospectively, as already exploited here for the decontamination of the support-plate motion, the platform is equally suited to multi-point monitoring of a single test rig or structure, an application scenario that a set of individually installed point sensors would make considerably more cumbersome and costly to set up.

A further direction for future development concerns the relationship between the present approach and Digital Image Correlation (DIC). DIC provides a dense, full-field

map of strain and displacement over the entire imaged surface, at the cost of a substantially heavier computational and calibration chain (speckle pattern application, stereo-camera setup, and per-frame correlation over the whole field of view). Template matching, by contrast, tracks only a small number of user-selected regions, which is sufficient whenever point-wise or few-region displacement histories are the quantity of interest, as in the present validation. For this class of applications, TMM can offer a substantially lighter computational alternative to DIC, at the cost of the full-field strain information that only DIC provides. A direct, quantitative comparison between the two approaches, in terms of both accuracy and computational cost on the same test rig, is identified here as a promising direction for future work.

5. Conclusions

This paper has presented the development and the laboratory assessment of a vision-based displacement monitoring platform, implemented in Python and OpenCV around the Template Matching Method, and benchmarked against a reference LVDT applied during experimental quasi-static cyclic tests on the SPEAD seismic dissipation device, using two structurally different targets. The tracking of a small ROI centred on and dominated by the 81 mm circular marker achieved a coefficient of determination $R^2 = 0.9975$, with $MAE = 0.8905$ mm and $RMSE = 1.0570$ mm, while tracking of a considerably larger ROI on the flange of an HEB 300 steel section, in which the same marker occupied only a minor fraction of the frame achieved $R^2 = 0.9932$, $MAE = 1.2214$ mm and $RMSE = 1.5133$ mm. In both configurations, the near-coincidence between MAE and RMSE indicates the absence of impulsive errors or tracking anomalies, and the mean percentage residuals, lower than 6% in both target cases, confirm the absence of any significant measurement drift effect. Importantly, the fact that near-millimetre accuracy was obtained even when tracking was dominated by the surrounding natural surface rather than by the marker itself indicates that the platform is not strictly dependent on a marker occupying a large share of the tracked frame, which is a favourable property for its future application to existing structural components where only a small marker, or none at all, can be applied.

These results demonstrate that a comparatively simple, camera-based measurement chain can reproduce the accuracy of a contact transducer to near-millimetre accuracy under controlled acquisition conditions, on both a small, marker-dominated ROI and a much larger ROI dominated by the natural structural surface. A further advantage that emerged clearly from the tests is that the camera observes the test apparatus as a whole rather than a single point, which made it possible to isolate and remove the rigid-body motion of the support plate directly from the recorded video, without any additional instrumentation, a capability that a conventional point sensor cannot offer. The platform therefore appears to be a credible, cost-effective and non-invasive complement to conventional contact instrumentation for laboratory testing and for the preliminary metrological qualification of vision-based monitoring procedures before their extension to real, in-service structures, where the acquisition distance, environmental disturbance and availability of suitable targets are expected to pose additional challenges. Rather than replacing established contact sensors, which retain unmatched metrological standards for permanent long-term monitoring, computer-vision-based displacement measurement should be regarded as an integrative tool, particularly valuable whenever rapid, low-cost and installation-free assessment of the static or dynamic structural response is required, including in post-earthquake emergency contexts where time and accessibility are at a premium.

Beyond the specific validation reported here, the processing platform itself constitutes a tangible research product of this study: a self-contained, general-purpose tool, requiring only a standard digital camera and a Python environment, that other researchers can adapt

to displacement-monitoring tasks on different structural elements and test setups without modification of its core tracking logic.

Author Contributions: Conceptualisation, V.P. and G.S.; methodology, V.P. and G.S.; software, V.P. and G.S.; validation, V.P. and G.S.; formal analysis, V.P. and G.S.; investigation, V.P. and G.S.; resources, V.P. and G.S.; data curation, V.P. and G.S.; writing—original draft preparation, V.P. and G.S.; writing—review and editing, V.P. and G.S.; visualisation, V.P. and G.S.; supervision, V.P. and G.S.; project administration, V.P. and G.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research benefited from equipment and experimental testing activities funded by the SAFER-REBUILT project (Sustainable Approaches For Earthquake Resistant Rehabilitation solutions for the BUILT environment), founded by the Extended Partnership Multi-Risk science for resilient communities under a changing climate (PE RETURN, <https://www.fondazionereturn.it/>)—Spoke VS3 Earthquake and Volcanoes (project code PE00000005, CUP H93C22000610002).

Data Availability Statement: The vision-based displacement measurement platform (including the graphical user interface) developed and used in this study, together with the data supporting the reported results, are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors have no relevant financial or non-financial interests to disclose.

References

1. Brownjohn, J.M.W.; Xu, Y.; Hester, D. Vision-based bridge deformation monitoring. *Front. Built Environ.* **2017**, *3*, 23. [\[CrossRef\]](#)
2. Santarsiero, G.; Picciano, V. Durability enhancement of half-joints in RC bridges through external prestressed tendons: The Musmeci Bridge's case study. *Case Stud. Constr. Mater.* **2023**, *18*, e01813. [\[CrossRef\]](#)
3. Micozzi, F.; Morici, M.; Zona, A.; Dall'Asta, A. Vision-based structural monitoring: Application to a medium-span post-tensioned concrete bridge under vehicular traffic. *Infrastructures* **2023**, *8*, 152. [\[CrossRef\]](#)
4. Feng, D.; Feng, M.Q.; Ozer, E.; Fukuda, Y. A vision-based sensor for noncontact structural displacement measurement. *Sensors* **2015**, *15*, 16557–16575. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Tonelli, D.; Caspani, V.F.; Valentini, A.; Rocca, A.; Torboli, R.; Vitti, A.; Perissin, D.; Zonta, D. Interpretation of bridge health monitoring data from satellite InSAR technology. *Remote. Sens.* **2023**, *15*, 5242. [\[CrossRef\]](#)
6. Bradski, G. The OpenCV library. *Dr. Dobbs's J.* **2000**, *120*, 122–125.
7. Dong, C.Z.; Catbas, F.N. A review of computer vision-based structural health monitoring at local and global levels. *Struct. Health Monit.* **2021**, *20*, 692–743. [\[CrossRef\]](#)
8. Civera, M.; Zanotti Fragonara, L.; Antonaci, P.; Anglani, G.; Surace, C. An experimental validation of phase-based motion magnification for structures with developing cracks and time-varying configurations. *Shock Vib.* **2021**, *2021*, 5518163. [\[CrossRef\]](#)
9. Guo, L.; Guo, W.; Chen, D.; Duan, B.; Shi, Z. Structural vibration measurement based on improved phase-based motion magnification and deep learning. *Mech. Syst. Signal Process.* **2025**, *224*, 111945. [\[CrossRef\]](#)
10. Santarsiero, G. Automated assessment of bridge guardrails for regional prioritization based on open-source data and deep learning algorithms. *Results Eng.* **2025**, *26*, 105210. [\[CrossRef\]](#)
11. Liu, Y.; Nie, X.; Fan, J.; Liu, X. Image-based crack assessment of bridge piers using unmanned aerial vehicles and three-dimensional scene reconstruction. *Comput.-Aided Civ. Infrastruct. Eng.* **2020**, *35*, 511–529. [\[CrossRef\]](#)
12. Masi, A.; Santarsiero, G.; Digrisolo, A.; Chiauzzi, L.; Manfredi, V. Procedures and experiences in the post-earthquake usability evaluation of ordinary buildings. *B. Geofis. Teor. Appl.* **2016**, *57*, 199–220. [\[CrossRef\]](#)
13. Masi, A.; Santarsiero, G.; Chiauzzi, L.; Gallipoli, M.R.; Piscitelli, S.; Vignola, L.; Bellanova, J.; Calamita, G.; Perrone, A.; Lizza, C.; et al. Different damage observed in the villages of Pescara del Tronto and Vezzano after the M6.0 August 24, 2016 Central Italy earthquake and site effects analysis. *Ann. Geophys* **2016**, *59*, Fast Track 5. [\[CrossRef\]](#)
14. Shi, H.; Zhang, M.; Jin, T.; Shi, X.; Zhang, J.; Xu, Y.; Guo, X.; Cai, X.; Peng, W. Computer vision-based monitoring of bridge structural vibration during incremental launching construction. *Buildings* **2025**, *15*, 1139. [\[CrossRef\]](#)
15. Ponsi, F.; Buoli, E.; Eslami Varzaneh, G.; Bassoli, E.; Briseghella, B.; Vincenzi, L. Vision-based approach for the static and dynamic monitoring of bridges. *Procedia Struct. Integr.* **2024**, *62*, 946–954. [\[CrossRef\]](#)
16. Tan, S.; Zhu, L.; Li, J.H.; Cui, Z.R.; Zhao, G.Y.; Tu, J.Y. Non-contact vibration displacement measurement method for railway bridges based on computer vision. *Structures* **2025**, *80*, 109803. [\[CrossRef\]](#)
17. Wei, X.; Wang, C.; Wang, X.; Du, Y. Structural health monitoring using urban surveillance cameras: Self-calibrated computer vision for enhanced displacement measurement. *Structures* **2025**, *75*, 108714. [\[CrossRef\]](#)

18. Lee, J.; Lee, K.C.; Cho, S.; Sim, S.H. Computer vision-based structural displacement measurement robust to light-induced image degradation for in-service bridges. *Sensors* **2017**, *17*, 2317. [[CrossRef](#)] [[PubMed](#)]
19. Santarsiero, G.; Manfredi, V.; Masi, A. Numerical evaluation of the steel plate energy absorption device (SPEAD) for seismic strengthening of RC frame structures. *Int. J. Civ. Eng.* **2020**, *18*, 835–850. [[CrossRef](#)]
20. Ielpo, P.; Santarsiero, G.; Picciano, V.; Masi, A.; Manfredi, V.; Ventura, G. Seismic retrofit of RC beam-column joints with external dissipation devices: A numerical study. In Proceedings of the 10th International Conference on Computational Methods in Structural Dynamics and Earthquake Engineering, Rhodes Island, Greece, 15–18 June 2025. [[CrossRef](#)]
21. Luan, L.; Liu, Y.; Sun, H. Extracting high-precision full-field displacement from videos via pixel matching and optical flow. *J. Sound Vib.* **2023**, *565*, 117904. [[CrossRef](#)]
22. Bendat, J.S.; Piersol, A.G. *Random Data: Analysis and Measurement Procedures*, 4th ed.; Wiley: Hoboken, NJ, USA, 2010.
23. Python Software Foundation. *Python*. 2026. version 3.13.5 [Computer software]. Available online: <https://www.python.org/> (accessed on 7 June 2026).
24. Xiong, B.; Zhang, Q.; Baltazart, V. On Quadratic Interpolation of Image Cross-Correlation for Subpixel Motion Extraction. *Sensors* **2022**, *22*, 1274. [[CrossRef](#)] [[PubMed](#)]
25. Marendić, A.; Gajski, D.; Duvnjak, I.; Kosor, A. Experimental accuracy evaluation of UAV-based homography for static and dynamic displacement monitoring of structures. *Sensors* **2026**, *26*, 1593. [[CrossRef](#)] [[PubMed](#)]
26. Marendić, A.; Gajski, D.; Duvnjak, I.; Paar, R. Evaluation of photogrammetric methods for displacement measurement during structural load testing. *Remote Sens.* **2025**, *17*, 2569. [[CrossRef](#)]
27. Wu, T.; Zhu, S.; Yuan, Q.; Zhu, W.; Lang, L.; Tang, H.; Wang, X. A moving window template matching method for vision-based full-field displacement measurement of bridges by tracking natural features. *Structures* **2026**, *90*, 112407. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.