16th World Congress in Computational Mechanics (WCCM)
21-26 July 2024, Vancouver, Canada
A. Korobenko, M. Laforest, S. Prudhomme, R. Vaziri (Eds)

COMPUTER VISION-BASED DYNAMIC DISPLACEMENT RECOGNITION METHOD FOR STRUCTURAL FRAMEWORKS

MUYANG HE[1], JIAYI HE[2] AND XIAODAN SUN[3]

[1]Civil Engineering College, Harbin Institute of Technology
73 Huanghe Road, Harbin, Heilongjiang, China
1308318090@qq.com

[2] School of Information Science and Engineering, Harbin Institute of Technology
2 Wenhuaxi Road, Weihai, Shandong, China
23S030127@stu.hit.edu.cn

[3]College of Aerospace and Civil Engineering, Harbin Engineering University
145 Nantong Street, Harbin, Heilongjiang, China
sunxiaodan@hrbeu.edu.cn

Key words: Computer vision, displacement recognition, scale factor, modal identification.

Abstract. Displacement induced by external forces is one of the most intuitive variables for assessing structural safety. Traditional contact methods, such as deploying displacement sensors on structures, are often limited by objective factors. Non-contact methods, such as utilizing computer vision algorithms like optical flow estimation and feature matching, offer the advantages of rapid and accurate structural displacement acquisition, unaffected by the structure itself and quick deployment. However, enhancing the accuracy of displacement monitoring based on computer vision remains a focal point in this field of study. In this paper, based on experimental data from a vibration table testing a four-layer reinforced concrete framework under three different conditions, we propose a method for processing dynamic displacement data that combines non-contact and contact approaches. This method integrates dynamically recognized structural displacements based on computer vision technology with data recorded by acceleration sensors on the structure to enhance displacement monitoring accuracy. The research results demonstrate that our method can obtain structural dynamic displacements based on computer vision information, confirming the effectiveness and reliability of our approach.

1 INTRODUCTION

With the rapid development of industries such as construction and transportation, the safety of infrastructure structures has garnered widespread attention. These structures inevitably experience damage such as corrosion and cracks during their service life. As time progresses, localized damage accumulates, leading to the gradual degradation of overall structural performance, thereby affecting safety and potentially causing significant accidents and casualties[1-3]. Effective health monitoring and damage diagnosis are necessary to ensure the normal use of these structures, enabling timely repairs and reinforcement during their service life[4]. Structural Health Monitoring (SHM)[5-7] finds extensive applications in civil engineering, with various types of infrastructure equipped with monitoring systems comprising diverse sensors. These systems generate substantial SHM data daily for users to analyze the status of buildings[8-13]. SHM has emerged as an increasingly popular research direction in the academic community of civil engineering[14]. It aims to monitor and analyze various loads, structural stresses, and strains reactions during the structure's usage, to assess its condition and predict its remaining service life, assisting managers in maintenance[15]. Among these, displacement resulting from structural deformations is one of the most crucial indicators for assessing structural health[16]. Currently, the civil engineering industry employs two main categories of methods, namely contact and non-contact, for structural displacement measurement[17-19].

Contact displacement measurement methods are primarily accomplished through sensors deployed on structures. Strain gauges are a common type of contact sensor, adhered to the structure's surface to measure strain. There exists a relationship between strain and displacement, allowing the derivation of displacement information by measuring strain[20]. Displacement transducers directly measure the structure's displacement, including linear displacement transducers, sliding potentiometers, and sliding inductors. They are typically mechanically coupled to the structure[21]. Accelerometers are devices used to measure the acceleration of objects in space. These sensors detect acceleration and output measurement results in digital or analog signals[22]. Accelerometers operate based on Newton's second law, often utilizing technologies such as microelectromechanical systems (MEMS), piezoelectric effects, or surface microelectromechanical systems (MEMS). MEMS accelerometers are the most common type, sensing acceleration changes through tiny mechanical structures typically made of silicon and integrated onto chips, resulting in compact, cost-effective sensors[23]. Accelerometers typically measure acceleration along three axes: X, Y, and Z, enabling capturing of objects' movements in three-dimensional space. Contact displacement measurement methods using sensors face challenges such as high costs and low measurement efficiency[24]. Additionally, traditional sensors pose difficulties in deployment in certain regions, making displacement measurement challenging through conventional sensor methods[4].

With the continuous advancement of computer vision technology and the increasing popularity of updated video capture devices, non-contact computer vision-based methods for dynamically identifying structural displacement are continuously evolving. Computer vision-based structural displacement measurement methods offer advantages such as low cost, non-contact operation, high precision, convenience, rapidity, and multi-point detection, making them increasingly prevalent in practical engineering applications [24-27]. In recent years, the development of non-contact visual displacement measurement systems has broadened new avenues for displacement measurement, primarily achieved through template matching/registration techniques [28]. Busca et al. [29] developed a visual displacement sensor system utilizing three template matching algorithms—pattern matching, edge detection, and digital image correlation (DIC). Visual sensors are employed to measure the vertical displacement of railway bridges by tracking high-contrast target panels fixed on the bridges. Song et al. [30] measured the displacement of cantilever beams from visual sensors by using sub-pixel Hough transform markers extracted from video images. Kim et al. [31] proposed a visual monitoring system using DIC to assess cable forces in cable-stayed bridges. Ribeiro et al. [32] utilized the RANdom SAmple Consensus (RANSAC) algorithm to measure dynamic displacement in railway bridges. The principle of computer vision-based methods involves processing video data collected by video capture devices, identifying structural features in digital images frame by frame, tracking and recognizing them, thereby calculating the motion trajectory of structural feature points in video frames, and consequently determining their pixel displacement [33]. The geometric relationship between the image frame and the real structure enables the calculation of a scaling factor, which transforms pixel displacement into actual structural displacement [28]. This method, combined with wireless communication technology, enables remote monitoring and automatic surveillance of structures.

This paper proposes a computer vision-based framework for dynamic displacement identification of structural frameworks, aiming to offer a novel approach for practical engineering applications in structural health monitoring. Leveraging experimental data from a vibration table test conducted on a four-story reinforced concrete framework under three different conditions, a hybrid non-contact and contact-based dynamic displacement data processing method is introduced. This method integrates computer vision technology-based structural dynamic displacement identification with data recorded by acceleration sensors mounted on the structure, thereby enhancing displacement monitoring accuracy. The experiments involve installing cameras at fixed points away from the test structure and capturing remote videos from fixed positions, eliminating the need for direct contact with the structure akin to displacement sensors. Additionally, the remote video coverage provides a larger measurement range, facilitating multi-point displacement measurements on the structure. In the visual recognition component, the widely used and operationally simple template matching method is employed for structural displacement measurement based on computer vision, owing to its early adoption and maturity in research both domestically and internationally. Therefore, in the proposed computer vision-based framework for dynamic displacement identification of structural frameworks, the template matching method is chosen for target image recognition and tracking [29]. Furthermore, the paper introduces the concept of regions of interest (ROI) to pre-segment video frames, aiming to reduce algorithmic time complexity.

2 VISUAL RECOGNITION

Figure 1 depicts a 1/2 scale four-story reinforced concrete frame structure constructed in the Civil Engineering Laboratory at Tongji University. The structural columns of this frame structure have a cross-section of 125 mm × 125 mm, while the beam sections measure 70 mm × 120 mm, with a floor thickness of 35 mm. The story height is 1500 mm, and on both sides of the frame structure columns, a cross-shaped visual target is drawn every 300 mm for computer vision-based structural dynamic displacement recognition.

The excitation source for the shake table testing was derived from the north-south components recorded at the EI Centro station during the 1940 Imperial Valley earthquake, the north-south components recorded at the Sylmar station during the 1994 Northridge earthquake, and a simulated Shanghai wave provided in the Shanghai Seismic Design Code. Various monitoring techniques were implemented during the shake table test, with sensor layouts as depicted in Figure 2. Absolute acceleration and relative displacement of each floor are typically measured using accelerometers and linear displacement sensors. A Sony camera is positioned 3.0 meters in front of the reinforced concrete frame, while a fixed camera is placed 8.0 meters in front of the frame.

Figure 1: 1/2 Scale Reinforced Concrete Frame Structure

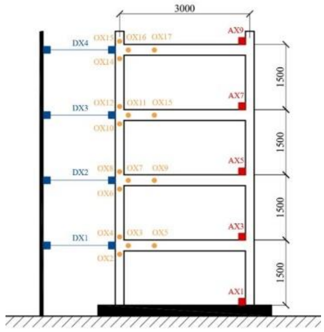


Figure 2: Layout of Sensors on the Shake Table

Within the video data captured by the visual acquisition system, the area occupied by the visual targets in each frame is relatively small compared to the entire image. Processing the entire image would significantly increase the algorithm's time complexity. Hence, it is advantageous to extract the region of interest (ROI) containing the visual targets as a small portion of the video frame. Subsequent algorithmic processing can then be focused solely on this ROI.

In the shake table experimental data of the reinforced concrete frame structure investigated in this study, visual targets were arranged as cross-shaped markers at intervals of 300 mm on both sides of the structure. Therefore, this study primarily focuses on the recognition of cross-shaped markers. Analysis of the video footage captured by the camera enables determination of the positions of these markers in each frame. In this study, the multi-target template matching method was employed to attempt recognition and tracking of the cross-shaped markers in the video, resulting in pixel displacements representing the structure's movements.

The recognition results of each frame of the test video are saved as individual text files. The relative pixel displacement from the 4th frame to the 1st frame is calculated using the distance formula between two points. As the dynamic displacement of the structure mainly manifests in the horizontal direction, with minimal displacement in the vertical and longitudinal directions, this study focuses solely on the horizontal displacement of the frame structure. Pixel displacements for each layer of the structure are computed, followed by upsampling and bandpass filtering to reduce errors. The processed pixel displacements are then converted into actual structural displacements. The overall procedure is illustrated in Figure 3.

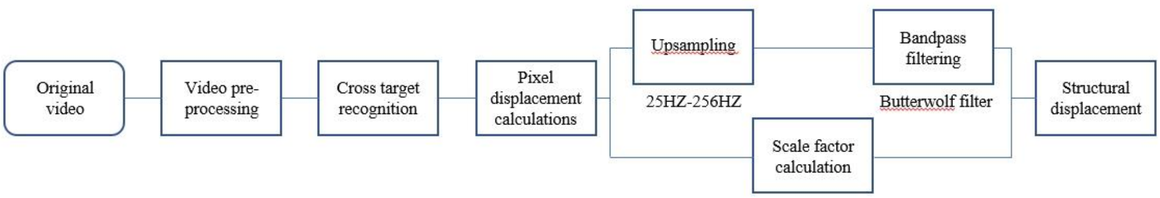


Figure 3 Illustrates the process of converting pixel displacement into actual displacement.

3 DATA FUSION

This study requires obtaining two sets of displacement data for the structure. The first set of displacement data consists of actual displacement values obtained after processing the pixel displacements of the crosshair targets extracted from the video using computer vision recognition technology with a scaling factor. The second set of displacement data comprises the dynamic displacements of the structure under vibration excitation obtained by integrating the acceleration of the structure measured by real-time accelerometers. By merging these two sets of data, we aim to simulate the real displacements of the structure under vibration excitation to the greatest extent possible. The specific process is illustrated in Figure 4.

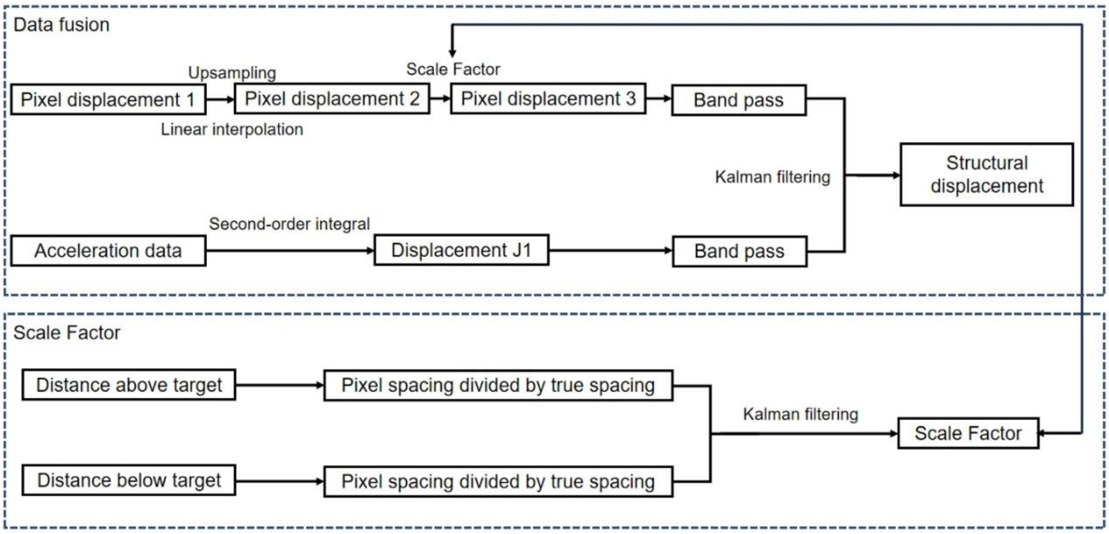


Figure 4 Flowchart of Data Fusion Method

To obtain the displacement of the structure under vibration excitation from acceleration signals, it is necessary to perform double integration on the discrete acceleration signals collected by the accelerometers. The frequency of the pixel displacement calculated through visual recognition is 25Hz, while the frequency of the displacement data obtained by double integrating the acceleration data collected by the accelerometer is also 256Hz. To unify the sampling frequency of the pixel displacement data and the displacement data obtained by double integrating acceleration, the upsampling factor between the two sets of data needs to be calculated, which represents the number of linearly interpolated values needed between every two data points of the pixel displacement. After calculation, the upsampling factor is determined to be 10.26.

The physical distance between the two crosshair targets on the structure is 300mm. By employing computer vision recognition methods, the pixel distance between the two crosshairs is determined. The scale factor for each point is calculated as 300mm divided by the pixel distance between the crosshairs. To reduce errors, the scale factor for the distance between the upper and lower segments of the crosshair is calculated frame by frame. The scale factors for both sets of data are fitted using a Kalman filter, and the fitted data are then averaged. The results of the scaling factor calculation for each layer are presented in Table 1.

Table 1: Scale Factor Caculation Results

Layer	Scale Factor mm/pix
1	3.4969
2	3.789963312
3	4.584388095
4	5.57358961

To attenuate the noise present in the structural vibration displacements obtained through visual recognition and those acquired via accelerometer sensors, this experiment employs Butterworth filters.

The filter order is set to 8, and the data is sampled at a frequency of 256Hz. Based on the characteristics of data from different operating conditions, distinct filter passband widths are required. After analysis, the passband for displacement ground truth data from the displacement sensor is set to 0.05Hz to 60Hz. For the first operating condition, the passband for acceleration data is set to 1.5Hz to 115Hz, for the second condition to 2.1Hz to 115Hz, and for the third condition to 1.5Hz to 115Hz.

For the fusion of the two sets of displacement data, Kalman filtering is employed. This involves merging the low-sampled displacement data with the high-sampled acceleration data. The position measured by the accelerometer serves as the state variable, while the position measured by the camera serves as the observation variable. The process involves updating, predicting, and updating steps, followed by visualization of the results.

4 CACULATION RESULTS

After collecting data from various experimental sensors, the proposed method in this study was employed for data processing, yielding the following experimental results as Figure5.

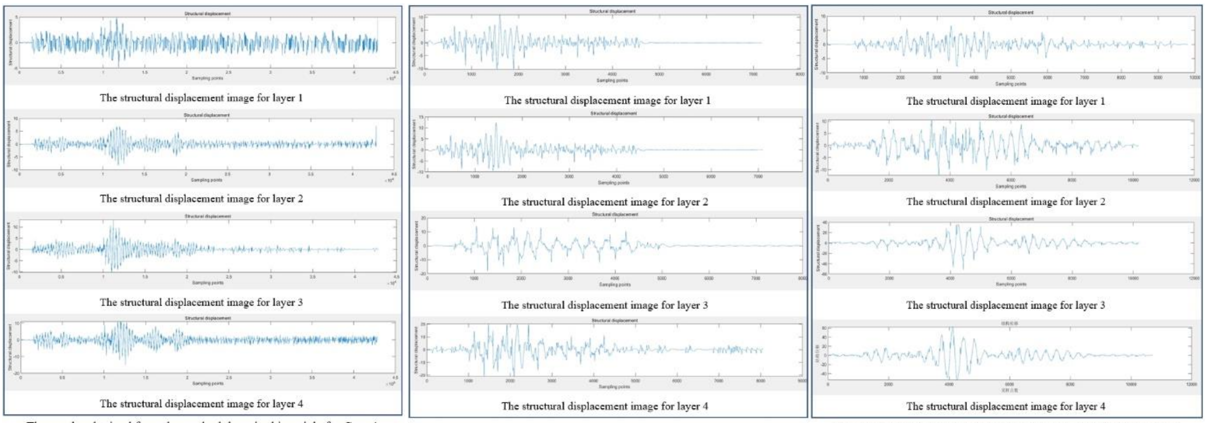
To validate the effectiveness and accuracy of the computer vision-based framework for dynamic displacement identification proposed in this paper, it is necessary to calculate the structural displacement estimation errors. Two metrics are employed for error calculation: peak displacement estimation error and root mean square (RMS) displacement estimation error. The structural displacement data obtained from displacement sensors are considered as the ground truth. A comparison is made between the data obtained from the proposed method and the data acquired from displacement sensors. Taking the first-floor displacement data of the framework structure obtained by displacement sensors under operating condition one as an example, the error between the two sets of data is computed using two methods:

Method 1: Direct comparison of the peak values of the two datasets.

The real values: max=7.3361mm， min=-5.7291mm; the values obtained from the method in this paper: max=6.2221mm， min=-4.8448mm.

The estimated peak values error: 7.3361-6.2221=1.114mm.

Method 2: Aligning the two datasets and subtracting them pointwise to calculate the error, then taking the maximum value as the peak error. Using this method, the peak displacement estimation error is calculated as 1.6195 mm, and the RMS displacement estimation error is calculated as 0.282847563.



The results obtained from the methodology in this article for Case 1 | The results obtained from the methodology in this article for Case 2 | The results obtained from the methodology in this article for Case 3

Figure 5 The results obtained from the methodology in this article

5 RESULT VALIDATION

The structural modal identification in this paper employs the covariance-driven stochastic subspace method. The core module of the program is a custom function FDDM(data, fs, FreqL, FreqU). The first parameter is the data under test, the second parameter is the sampling frequency of the data (128Hz for displacement ground truth and 256Hz for acceleration), and the third and fourth parameters are the lower and upper limits of the passband frequency, set to match the cutoff frequencies of the bandpass filter.

By inputting the data, this program can generate modal frequency distribution plots for each operating condition and each floor, with the horizontal axis corresponding to the modal frequencies. The true range of modal frequencies for the structure is calculated as 0 to 20.4984Hz, while the range obtained using the proposed method is calculated as 0 to 21.3825Hz, resulting in a 4.3% error. These results demonstrate the feasibility of the proposed method.

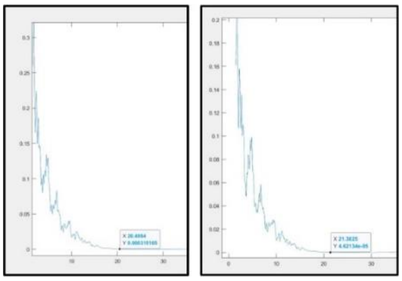


Figure 6 Comparison of Modal Parameters Obtained by the Proposed Method and Ground Truth

6 CONCLUSIONS

1.Through experimentation, it has been confirmed that employing the multi-target template matching method for visual recognition can achieve target capture and dynamic displacement output. The calculated peak displacement estimation error of 1.6195 mm and RMS displacement estimation error of 0.282847563 meet the practical requirements of building displacement monitoring, demonstrating engineering practicality and promotion value.

2.Compared to traditional methods using sensors for displacement monitoring, the proposed method does not require significant human and material resources for sensor manufacturing, installation, and fixation, making it more environmentally friendly. Computer vision technology evolves rapidly and is more advanced, making computer vision-based structural dynamic displacement identification methods sustainable.

3.The proposed algorithm offers convenient data acquisition and high precision in processing results, providing a basis for further research on measurement methods for small displacement vibrations in framework structures. This method presents a new and effective approach for building displacement monitoring, with significant practical value and market prospects.

4.The method in this paper involves manual selection of the region of interest (ROI). Future research will focus on automatic ROI selection methods to achieve automatic ROI selection.

REFERENCES

[1] Tumrate, C.S., Saini, D.K., Gupta, P. *et al.* Evolutionary Computation Modelling for Structural Health Monitoring of Critical Infrastructure. *Arch Computat Methods Eng* **30**, 1479–1493 (2023). https://doi.org/10.1007/s11831-022-09845-1

[2] Deep-learning-based crack detection with applications for the structural health monitoring of gas turbines.Mahtab Mohtasham Khani, Sahand Vahidnia, Leila Ghasemzadeh, Y Eren Ozturk, Mustafa Yuvalaklioglu, Selim Akin, Nazim Kemal Ure. Structural Health Monitoring, Vol 19, Issue 5, pp. 1440 – 1452. Issue published date: September-01-2020, 10.1177/1475921719883202

[3] Azimi, M.; Eslamlou, A.D.; Pekcan, G. Data-Driven Structural Health Monitoring and Damage Detection through Deep Learning: State-of-the-Art Review. *Sensors* **2020**, *20*, 2778. https://doi.org/10.3390/s20102778

[4] Sen Wang, Chang Liu, Yinhui Zhang, Fully convolution network architecture for steel-beam crack detection in fast-stitching images, Mechanical Systems and Signal Processing.

[5] Inman DJ, Farrar CR, Vicente LJ, et al. *Damage prognosis: for aerospace, civil, and mechanical systems. Chichester: John Wiley & Sons, 2005.*

[6] Sohn H, Farrar CR, Hemez FM, et al. A literature review on structural health monitoring: 1996-2001. Los Alamos, New Mexico: Los Alamos National Laboratory, 2004.

[7] Li SL, Li H, Liu Y, et al. Benchmark problem for SMC structural health monitoring using actual cable-stayed bridge monitoring data. Struct Control Health Monit 2014;21: 156–172.

[8] Li H, Ou JP. Recent advances in structural health monitoring of cable-stayed bridges. J Civ Struct Health Monit 2016;6: 43–67.

[9] Li H, Ou JP, Zhang XG, et al. Research and practice of health monitoring for large-span

bridges in mainland China. Smart Struct Syst 2015;15: 555–576.
[10] Ni YQ, Wong KY, Xia Y. Health monitoring from landmark bridges to towering structures. Adv Struct Eng 2011;14: 103–119.
[11] Dalvi A, Norouzi M, Hunt V, et al. Structural health monitoring system for the Ironton-Russell Bridge during substructure construction. In: SPIE 9805, Smart Structures and Materials + Nondestructive Evaluation and Health Monitoring, Las Vegas, Nevada, March 20-24, 2016. Bellingham, Washington: SPIE.
[12] Huang L, Duan W, Zhu J, et al. Deformation monitoring research during deep foundation construction of subway stations. Appl Mech Mater 2016;847: 425–430.
[13] Park JW, Kim K, Cho YK. Automation of construction safety monitoring framework using cloud-supported BIM and BLE mobile tracking sensors. J Constr Eng Manag 2016;143: 05016019.
[14] Islam, M.M.M.; Kim, J.-M. Vision-Based Autonomous Crack Detection of Concrete Structures Using a Fully Convolutional Encoder–Decoder Network. *Sensors* **2019**, *19*, 4251. https://doi.org/10.3390/s19194251
[15] Wang, L.; Li, J.; Kang, F. Crack Location and Degree Detection Method Based on YOLOX Model. Appl. Sci. 2022, 12, 12572. https://doi.org/10.3390/app122412572
[16] D. Ribeiro, R. Calçada, J. Ferreira, T. Martins, Non-contact measurement of the dynamic displacement of railway bridges using an advanced video-based system, Engineering Structures, Volume 75, 2014, Pages 164-180, ISSN 0141-0296, https://doi.org/10.1016/j.engstruct.2014.04.051.
[17] J. Lee, S. Cho, M. Shinozuka, C. Yun, C. Lee, W. Lee, Evaluation of bridge load carrying capacity based on dynamic displacement measurement using real-time image processing technique, Steel Struct, 6 (2006), pp. 377-385
[18] J. Lee, Y. Fukuda, M. Shinozuka, S. Cho, C. Yun, Development and application of a vision-based displacement measurement system for structural health monitoring of civil structures, Smart Struct Syst, 3 (2007), pp. 373-384
[19] Lee J, Oh J, Park M, Kwon S, Kwark J. Bridge displacement measurement system using image processing. IABMAS'06 – Bridge maintenance, safety, management, resilience and sustainability. Porto, Portugal; 2006.
[20] Chatterjee, K., Mahato, S.N., Chattopadhyay, S. et al. High accuracy displacement measuring system using strain gauge based displacement sensor and direct sequence spread spectrum techniques in data acquisition system. Instrum Exp Tech 60, 154–157 (2017). https://doi.org/10.1134/S0020441217010055
[21] Zhang, Z.; Wang, L.; Cao, B.; Zhang, H.; Liu, J. A Moving Magnetic Grid-Type Long-Range Linear Absolute Displacement Sensor. Sensors 2023, 23, 700. https://doi.org/10.3390/s23020700
[22] R. Liu, M. Liu, X. Sun and Y. Wei, "Signal Processing and Accelerometer-based Design for Portable Small Displacement Measurement Device," 2008 International Conference on Embedded Software and Systems, Chengdu, China, 2008, pp. 575-579, doi: 10.1109/ICESS.2008.65.
[23] S. Morichika, H. Sekiya, Y. Zhu, S. Hirano and O. Maruyama, "Estimation of Displacement Response in Steel Plate Girder Bridge Using a Single MEMS Accelerometer," in IEEE Sensors Journal, vol. 21, no. 6, pp. 8204-8208, 15 March15, 2021, doi: 10.1109/JSEN.2021.3051697.

[24] Alipour M, Harris DK, Miller GR (2019) Robust pixel-level crack detection using deep fully convolutional neural networks. Journal of Computing in Civil Engineering 33(6):04019040
[25] Andy Nguyen, Vahidreza Gharehbaghi, Ngoc Thach Le, Lucinda Sterling, Umar Inayat Chaudhry, Shane Crawford, ASR crack identification in bridges using deep learning and texture analysis, Structures, Volume 50, 2023, Pages 494-507, ISSN 2352-0124, https://doi.org/10.1016/j.istruc.2023.02.042.
[26] D. Feng, M.Q. Feng, Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection–A review, Eng Struct, 156 (2018), pp. 105-117
[27] A review of computer vision–based structural health monitoring at local and global levels. Chuan-Zhi Dong, F Necati Catbas. Structural Health Monitoring. Vol 20, Issue 2, pp. 692 – 743. Issue published date: March-01-2021. 10.1177/1475921720935585
[28] Feng, D.; Feng, M.Q.; Ozer, E.; Fukuda, Y. A Vision-Based Sensor for Noncontact Structural Displacement Measurement. Sensors 2015, 15, 16557-16575. https://doi.org/10.3390/s150716557
[29] Busca, G., Cigada, A., Mazzoleni, P. et al. Vibration Monitoring of Multiple Bridge Points by Means of a Unique Vision-Based Measuring System. Exp Mech 54, 255–271 (2014). https://doi.org/10.1007/s11340-013-9784-8
[30] Song Y-Z, Bowen CR, Kim AH, Nassehi A, Padget J, Gathercole N. Virtual visual sensors and their application in structural health monitoring. Structural Health Monitoring. 2014;13(3):251-264. doi:10.1177/1475921714522841
[31] Kim S-W, Jeon B-G, Kim N-S, Park J-C. Vision-based monitoring system for evaluating cable tensile forces on a cable-stayed bridge. Structural Health Monitoring. 2013;12(5-6):440-456. doi:10.1177/1475921713500513
[32] D. Ribeiro, R. Calçada, J. Ferreira, T. Martins, Non-contact measurement of the dynamic displacement of railway bridges using an advanced video-based system, Engineering Structures, Volume 75, 2014, Pages 164-180, ISSN 0141-0296, https://doi.org/10.1016/j.engstruct.2014.04.051.
[33] Zhuang, Y.; Chen, W.; Jin, T.; Chen, B.; Zhang, H.; Zhang, W. A Review of Computer Vision-Based Structural Deformation Monitoring in Field Environments. Sensors 2022, 22, 3789. https://doi.org/10.3390/s22103789