J. Electrical Systems 22-01 (2026): 376-384

[1]**Ahmed Alzughaibi**

Pixel-Domain Modal-Parameter-Based Structural Health Monitoring Using In-Building Cameras: A Review

Abstract: - Recent studies have shown that existing in-building surveillance cameras can be repurposed as practical sensors for structural health monitoring of buildings. While many vision-based structural monitoring methods aim to recover physical displacements, a smaller but important line of research has shown that useful structural information can also be extracted directly in the pixel domain. This review examines modal-parameter-based structural health monitoring approaches that use cameras installed inside buildings. The reviewed studies are organized into four stages in the development of the topic: proof-of-concept frequency extraction from existing indoor cameras, pixel-based operating-mode identification, synchronized multi-camera modal identification, and post-earthquake damage assessment through frequency-shift analysis. The review highlights the advantages of pixel-domain processing for practical deployment, especially the reduced calibration needs and compatibility with pre-installed cameras, while also discussing key limitations related to synchronization, camera motion, lighting, and the interpretation of pixel-domain outputs. The paper concludes by identifying promising directions for future investigation.

Keywords: Smart buildings, Smart cities, structural health monitoring (SHM), Vision-based monitoring, Disaster resilience

I. INTRODUCTION

Earthquakes can leave buildings with serious structural damage, and a considerable portion of that damage may not be apparent from an exterior inspection alone. As a result, post-earthquake decisions about safety and occupancy depend on obtaining reliable information about structural condition within a short time window. Conventional structural health monitoring (SHM) systems address this need through dedicated sensing devices, such as accelerometers, strain gauges, and displacement transducers, together with the associated communication and data-processing infrastructure [1-8]. Although these systems can deliver accurate measurements, their broader use in residential and commercial buildings is still restricted by installation cost, networking requirements, and maintenance needs [9-12].

In response to these limitations, vision-based SHM has attracted growing interest as a practical alternative. Cameras offer a non-contact and relatively low-cost means of measuring structural response directly from images or video recordings [13-17]. During the last decade, image-based monitoring methods have been studied for different classes of civil infrastructure, including bridges, buildings, and dams [18-23]. For buildings in particular, camera-based approaches have been used both for detecting visible damage and for measuring structural response [24-29]. However, many of these methods still depend on cameras deployed specifically for measurement, whether placed off the structure or mounted directly on it, which reduces their scalability in residential and commercial settings [22], [26], [30].

A more practical opportunity comes from the fact that many buildings already contain surveillance cameras that are installed for routine security and operational use, as illustrated in Figure 1. Because this infrastructure is already powered, networked, and maintained, it offers a potentially scalable foundation for structural monitoring. Indoor surveillance cameras are especially promising because they operate in a relatively controlled environment and often capture stable visual features on the floor below. Recent studies have shown that these cameras can be used not only for drift-related measurements, but also for extracting modal information, including dominant frequencies, operating modes, and frequency-based damage indicators [31-34].

It is useful to distinguish between deformation-based monitoring and modal-parameter-based monitoring. Methods aimed at drift estimation usually require image measurements to be converted into physical units, since the quantity of interest is displacement. In contrast, modal-parameter-based methods can often work directly with pixel-domain motion signals because the dominant frequency content is unaffected by amplitude scaling [31], [32]. This characteristic makes pixel-domain analysis particularly appealing for repurposed indoor surveillance

---

[1] Department of Electrical Engineering, College of Engineering, Qassim University, Saudi Arabia
*Corresponding author: Ahmed Alzughaibi ah.alzughaibi@qu.edu.sa

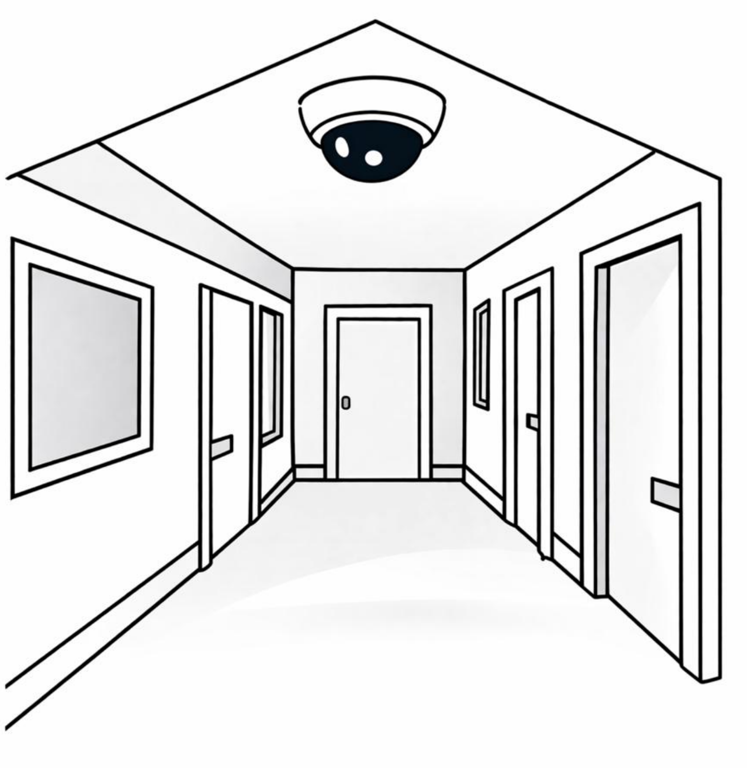

Figure 1. Conceptual illustration of an indoor surveillance-camera configuration.

cameras, where full calibration may be unavailable or impractical for broader deployment.

A number of review articles have addressed computer-vision-based SHM in a general sense [35-39], whereas others have concentrated on particular structural types or sensing platforms [40-42]. A focused review of pixel-domain modal-parameter-based monitoring using in-building surveillance cameras remains worthwhile, because this topic lies at the intersection of practical deployment constraints and vibration-based structural assessment.

Accordingly, this paper presents a review of modal-parameter-based SHM using in-building cameras, with emphasis on pixel-domain approaches. The development of this topic is discussed through four representative stages: early proof-of-concept studies on frequency extraction using indoor cameras [31], pixel-based operating-mode identification without conversion to physical units [32], synchronized multi-camera modal identification [33], and natural-frequency-shift-based seismic damage assessment [34].

The remainder of this paper is organized as follows. Section 2 discusses the motivation for pixel-domain modal monitoring using indoor cameras. Section 3 reviews the development of pixel-domain modal monitoring using indoor cameras. Section 4 provides a comparative discussion of the main contributions and limitations of the reviewed studies. Section 5 outlines future research directions. Section 6 concludes the paper.

II. MOTIVATION FOR PIXEL-DOMAIN MODAL MONITORING WITH INDOOR CAMERAS

The appeal of repurposing indoor surveillance cameras for SHM lies mainly in the possibility of reducing the cost and logistical challenges associated with dedicated sensing systems. Instead of installing a new monitoring network at each site, this approach takes advantage of hardware that is already embedded in the building environment. For residential and commercial buildings that are unlikely to be instrumented with conventional SHM systems, indoor cameras may therefore provide a practical additional source of structural information [7], [8], [10].

In many vision-based displacement methods, meaningful structural interpretation depends on mapping image-plane motion into physical units using a scale factor, planar homography, or a more complete camera model [30], [43], [44]. Although these steps are necessary when the goal is drift estimation, they can become impractical in surveillance-camera-based deployment because the existing viewpoints, scene geometry, and available in-scene references are often not ideal for accurate metric reconstruction. Modal-parameter-based monitoring offers a different advantage: dominant frequencies can often be identified directly from image-derived motion histories without requiring full pixel-to-metric conversion [31], [32].

Another benefit of modal-parameter-based analysis is that it can still provide useful structural information when deformation-based quantities are more difficult to recover. Since natural frequency is related to structural stiffness, changes in frequency can serve as damage-sensitive indicators [34]. In addition, when multiple camera views are synchronized, their combined measurements can be used to estimate mode-related response and, in some cases, reconstruct mode shapes [32], [33].

At the same time, pixel-domain approaches also face practical limitations. Reliable frequency estimation still depends on robust feature tracking and sufficient video quality. Multi-camera implementations introduce synchronization requirements, and pixel-based operating modes do not have exactly the same interpretation as conventional physical mode shapes [32], [33]. Their practical performance may also be influenced by the suitability of the scene itself, including lighting conditions, motion blur, and the presence of nonstructural objects within the field of view, all of which remain recurring concerns in indoor-camera-based structural monitoring [45-47].

III. DEVELOPMENT OF INDOOR CAMERA PIXEL-DOMAIN MODAL MONITORING

A. *Early proof of concept using existing indoor cameras*

Harvey and Elisha [31] presented one of the earliest demonstrations that existing cameras installed within a building can be used for vibration monitoring. Their study showed that indoor surveillance-style cameras could be used to extract inter-story drift and, from it, dynamic characteristics such as the fundamental period. The work was important because it shifted the discussion from installing a camera for monitoring to leveraging cameras that are already present in the building environment.

Their image-processing workflow relied on a straightforward target-based procedure. A region of interest was selected, thresholding was applied to identify the target, and connected components were tracked frame by frame. The study also discussed temporal aliasing as a practical issue for low-frame-rate surveillance video and noted that tracking ambiguity becomes more severe when large displacements occur between frames. Although this work still included pixel-to-physical conversion, it established a crucial point for later studies: indoor surveillance footage can contain structurally meaningful vibration information even when the camera arrangement is not designed for traditional structural monitoring [31].

The reported results were especially important for the later pixel-domain direction. In the lab-scale validation, the flexible camera supports had their own identified frequencies of 27.2, 12.2, and 5.16 Hz for cantilevers A, B, and C, respectively, yet the building frequencies remained clearly identifiable. From the video-based and accelerometer-based spectra, the first three building frequencies were observed at approximately 3.1, 9.1, and 13.4 Hz, with excellent agreement between the rigid-camera reference, the flexible-support surveillance camera, and the accelerometer measurements [31]. The study therefore showed that while flexible support degraded displacement accuracy, it did not significantly affect natural-frequency identification. Another important observation from this early work is that the vibration information remains useful even when the camera support is not ideally rigid, which helped motivate later efforts to focus more directly on modal quantities rather than only on physical displacement reconstruction.

B. *Pixel-based operating modes without unit conversion*

A major conceptual step was introduced by Hosseinzadeh and Harvey [32], who proposed the use of pixel-based operating modes derived directly from synchronized surveillance videos. In this study, robust feature points were identified automatically and tracked using the Kanade-Lucas-Tomasi (KLT) algorithm. The resulting relative pixel motions from multiple cameras were transformed into the frequency domain to estimate dominant structural frequencies and operating-mode-like quantities directly in pixel units.

The significance of this contribution lies in the removal of the intermediate calibration step. Rather than converting image motion into engineering units before structural interpretation, the method treated the relative pixel motions themselves as the basis for modal analysis. The identified pixel-based operating modes do not have the same physical interpretation as conventional mode shapes, but they provide repeatable signatures of the structural dynamic behavior and can be compared across tests to detect changes [32].

The reported results showed that the method was both stable and sensitive to structural change. For the reference configuration (A), the mean estimated frequencies were about 3.26, 9.82, and 14.36 Hz, whereas for the modified configuration (B) they were about 3.22, 9.03, and 13.97 Hz, showing a consistent frequency reduction after the physical modification to the structure [32]. In addition, the extracted pixel-based operating modes were highly repeatable, with MAC values at or above 0.90 across repeated tests. When the pixel-based modes of configuration A were compared with those of configuration B, the MAC values decreased in all modes, indicating that the method could detect the presence of structural change; among the three modes, the third mode was reported to be the most sensitive to the modification. This work therefore demonstrated that target-free processing is feasible for indoor modal monitoring, provided that the recorded scene contains sufficiently strong and trackable features, and in that sense, it moved the field closer to practical surveillance-camera repurposing by reducing dependence on specially installed targets and on full geometric calibration.

C. Synchronized multi-camera modal identification

The next important development came from the extension of indoor surveillance-camera monitoring to synchronized multi-story modal identification. Hosseinzadeh et al. [33] proposed a framework that uses multiple interior cameras to reconstruct a more complete representation of building motion. In this study, cameras at different stories measured inter-story drifts and rotations by tracking the motion of the floor below, and the resulting measurements were synchronized and processed using frequency domain decomposition.

This work is particularly important because it bridges pixel-domain ideas and more conventional modal identification. Unlike the earlier pixel-only approach, the study used the synchronized measurements to identify modal frequencies and mode shapes [33]. The framework therefore showed that interior surveillance cameras are not limited to frequency extraction; with proper synchronization and processing, they can also support multi-degree-of-freedom modal analysis.

The reported results showed strong agreement with accelerometer-based identification. The first mode was identified most accurately, with MAC values very close to unity in nearly all cases, while the second and third modes were recovered with lower but still acceptable agreement [33]. In general, acceptable MAC values greater than about 0.8 were reported in almost all cases for excitation levels of 5%g RMS or greater, and the first mode was recovered especially well because its contribution to the displacement responses was dominant [33]. In the repeatability assessment, the vision-based extracted modes remained consistent across tests, with MAC values of 0.76-0.99 for the higher modes depending on excitation level, and values consistently above about 0.95 at excitation levels greater than 5%g RMS. Moreover, when the modified structural configuration was compared with the baseline case, the average MAC values for the first, second, and third modes were about 0.96, 0.87, and 0.79, respectively, indicating that the method was also sensitive to physical changes in the structure. The study also brought attention to practical synchronization issues. When multiple cameras are used, differences in start time and slight deviations between nominal and actual frame rate can affect phase relationships and therefore influence modal identification. This is a central issue for any future multi-camera deployment in real buildings.

D. Frequency-shift-based post-earthquake damage assessment

A more recent step in the literature is the move from modal identification to damage assessment using natural-frequency-based indicators extracted from indoor surveillance videos. Zhou et al. [34] proposed a strategy that combines computer vision with time-frequency analysis to estimate the structural natural frequency from inter-story pixel displacements and then compute a damage index from the resulting frequency shift.

This contribution is important for two reasons. First, it directly addresses a major practical challenge in post-earthquake use of surveillance cameras: it is often difficult to recover calibrated displacement histories or synchronized multi-camera mode shapes immediately after an event. Second, it shows that a frequency-based damage indicator may still be obtained from a single video sequence, which reduces the burden on instrumentation and data collection [34].

The study used marker-free feature tracking together with time-frequency analysis, specifically a synchroextracting-transform-enhanced empirical wavelet transform (SET-EWT), to estimate the time-varying natural frequency [34]. The reported results were promising in both the small-scale and full-scale validations. In the small-scale verification, the maximum error in the estimated damage index was 0.097, and the damage indices estimated from first-story and second-story responses were consistent, indicating that the approach could obtain a natural-frequency-based damage index using only one camera [34]. In the full-scale reinforced-concrete benchmark model, the first natural frequency changed from 0.84 Hz before the earthquake to 0.62 Hz after the earthquake, which corresponded to a damage index of 0.455 and was classified as moderate damage. By doing so, the paper extended the indoor-camera monitoring concept from modal characterization to direct seismic damage assessment and showed that image-derived frequency shifts can support practical post-earthquake screening [8], [10], [34].

Although all four studies focus on extracting modal information from indoor camera recordings, they differ in their monitoring objectives, tracking strategies, signal-processing pipelines, and final structural outputs. To facilitate a clearer comparison, Table 1 summarizes the main characteristics of these studies, including the camera configuration, validation setting, and the specific contribution of each work to the development of indoor pixel-domain modal monitoring. The typical processing workflow followed by these methods is illustrated in Figure 2, where the analysis proceeds from camera recording and pre-processing to motion extraction, modal identification, and finally damage-sensitive indicators.

Table 1. Comparison of representative modal-parameter-based structural health monitoring studies using in-building surveillance cameras.

Category	**Harvey and Elisha (2018)** [31]	**Hosseinzadeh and Harvey (2019)** [32]	**Hosseinzadeh et al. (2021)** [33]	**Zhou et al. (2024)** [34]
Main objective	Natural frequency estimation under flexible camera support	Pixel-only modal identification and pixel-based operating modes	Modal identification and 3D motion of floors using synchronized cameras	Seismic damage assessment using natural-frequency shift
Tracking strategy	Target-based dot-grid centroid tracking	Target-free KLT feature tracking	Target-based tracking with synchronized multi-camera views	Target-free KLT feature tracking
Signal-processing method	Pixel-based PSD / FFT periodogram	Pixel-domain FFT and pixel-based operating modes	Modal identification using synchronized multi-camera measurements	Instantaneous frequency analysis using SET-EWT
Main structural output	Natural frequency	Modal frequencies and pixel-based operating modes	Modal frequencies, mode shapes, and torsional behavior	Frequency-shift-based damage index
Camera / setup characteristics	Indoor camera with flexible support; experimental validation in lab-scale and full-scale shake-table tests	Multiple indoor cameras observing a three-story lab structure on a shake table	Synchronized multi-camera indoor network for multi-story response measurement	Indoor surveillance-style camera applied to small-scale and full-scale building tests
Validation type	Experimental: lab and full-scale shake table	Experimental: three-story lab structure on shake table	Numerical and experimental validation	Experimental: small-scale and full-scale building tests
Main contribution	Showed that natural-frequency estimates remain reliable even with flexible camera support	Proposed pixel-based operating modes for SHM without physical-unit conversion	Extended indoor-camera monitoring to synchronized multi-camera modal identification	Linked indoor-camera measurements to seismic damage assessment through frequency-shift analysis
Main relevance to this review	Established early proof of concept for extracting vibration information from indoor cameras	Introduced the core idea of modal-parameter extraction directly in the pixel domain	Showed how indoor camera networks can support richer system-level modal identification	Demonstrated how pixel-domain or image-derived frequency shifts can support damage assessment

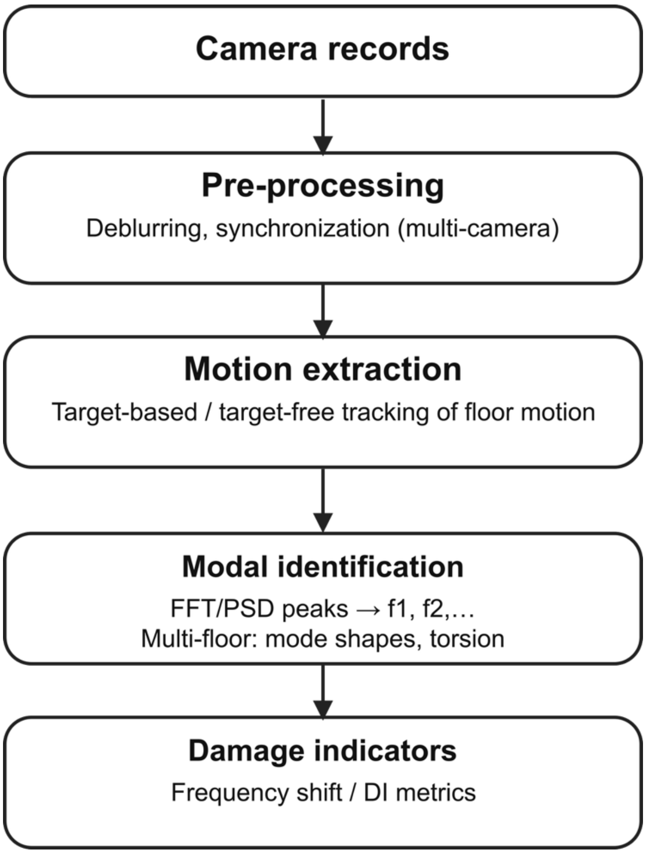

Figure 2. Typical processing flow for indoor camera based modal-parameter structural health monitoring.

IV. DISCUSSION

The reviewed studies show a clear progression in the development of indoor surveillance-camera-based modal monitoring. Harvey and Elisha [31] established the feasibility of extracting structural dynamic information from existing indoor cameras. Hosseinzadeh and Harvey [32] then demonstrated that modal information could be extracted directly in pixel units, thereby reducing dependence on physical calibration. Hosseinzadeh et al. [33] extended the concept to synchronized multi-camera modal identification, while Zhou et al. [34] connected frequency-based measurements to seismic damage assessment.

One of the main strengths of pixel-domain modal monitoring is the reduced calibration burden. Since dominant frequencies are preserved under amplitude scaling, frequency estimation does not require pixel-to-millimeter conversion [31], [32]. This makes the approach well suited to existing indoor surveillance systems, where geometry and calibration references may be uncertain or unavailable. Another advantage is that multi-camera pixel-domain measurements can provide operating-mode-like information, and in more advanced configurations they can support modal identification and damage-sensitive analysis [32-34].

At the same time, the literature also reveals important limitations. First, tracking performance depends on scene texture, visibility, lighting conditions, and image quality. These issues are not unique to indoor monitoring, but they become particularly important when the cameras were installed for surveillance rather than measurement [16], [29], [45], [46]. Second, multi-camera approaches remain sensitive to synchronization errors [33]. Third, while pixel-based operating modes are valuable indicators, they are not identical to conventional physical mode shapes, which means their interpretation must be made carefully [32]. Finally, frequency-shift-based damage indicators are promising, but additional validation under more realistic building conditions is still needed [34].

Overall, the literature suggests that modal-parameter-based monitoring should be viewed as complementary to deformation-based monitoring. Drift-based measures remain more directly connected to post-earthquake performance limits, whereas pixel-domain modal quantities are often easier to obtain from pre-installed indoor cameras [31], [34], [43], [44]. This complementarity is one of the main reasons why pixel-domain modal monitoring deserves focused attention.

V. FUTURE RESEARCH DIRECTIONS

Several research directions appear especially important for the next stage of development.

First, there is a need for more realistic validation under conditions that resemble actual post-earthquake operation. Many current demonstrations were performed in controlled laboratory settings with favorable image quality. Future work should examine how frequency estimation and pixel-based modal outputs behave under illumination loss, dust, smoke, partial occlusion, and scene changes.

Second, synchronization deserves further attention. Multi-camera indoor monitoring is one of the most promising directions because it can move the field beyond frequency extraction toward richer system identification [33]. However, the effect of timing errors, frame-rate mismatch, and dropped frames on identified modal quantities still needs more systematic study.

Finally, future work should consider deployment logic at the building-network level. The long-term value of indoor surveillance-camera monitoring is not only methodological but also operational. Event-triggered video retrieval, automated processing, privacy-preserving workflows, and rapid building-level tagging will likely be central to any practical implementation [35], [37], [38].

VI. CONCLUSIONS

This paper presented a focused review of modal-parameter-based structural health monitoring using in-building surveillance cameras, with particular attention to pixel-domain processing. The reviewed studies show a clear development path. Early work established that existing indoor cameras can recover structurally meaningful vibration information. Later studies demonstrated that dominant frequencies and operating-mode-like quantities can be extracted directly from relative pixel motions without full pixel-to-metric conversion. Subsequent work extended the concept to synchronized multi-camera modal identification, and more recent studies linked indoor-camera measurements to post-earthquake damage assessment through frequency-shift-based indicators [31-34].

These studies indicate that pixel-domain modal monitoring is a promising direction for practical and scalable building assessment using pre-installed surveillance infrastructure. Its main strengths are the reduced calibration burden and compatibility with existing camera systems, while its main challenges include synchronization, tracking robustness, and interpretation of pixel-domain outputs. Continued progress in these areas may help move indoor-camera-based monitoring from controlled demonstrations toward broader post-earthquake structural assessment applications.

ACKNOWLEDGMENT

The Researchers would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for financial support QU-APC-2026.

REFERENCES

[1] D.-H. Kim and M. Q. Feng, “Real-Time Structural Health Monitoring Using a Novel Fiber-Optic Accelerometer System,” *IEEE Sensors Journal*, vol. 7, no. 4, pp. 536–543, Apr. 2007, doi: 10.1109/JSEN.2007.891988.

[2] P. Ragam and N. Devidas Sahebraoji, “Application of MEMS-based accelerometer wireless sensor systems for monitoring of blast-induced ground vibration and structural health: A review,” *IET Wireless Sensor Systems*, vol. 9, no. 3, pp. 103–109, Jun. 2019, doi: 10.1049/iet-wss.2018.5099.

[3] L. Zhu, Y. Fu, R. Chow, B. F. Spencer, J. W. Park, and K. Mechitov, “Development of a High-Sensitivity Wireless Accelerometer for Structural Health Monitoring,” *Sensors*, vol. 18, no. 1, p. 262, Jan. 2018, doi: 10.3390/s18010262.

[4] G. Liu *et al.*, “Review of Wireless RFID Strain Sensing Technology in Structural Health Monitoring,” *Sensors*, vol. 23, no. 15, p. 6925, Jan. 2023, doi: 10.3390/s23156925.

[5] A. Ghosh, D. J. Edwards, M. R. Hosseini, R. Al-Ameri, J. Abawajy, and W. D. Thwala, “Real-time structural health monitoring for concrete beams: A cost-effective ‘Industry 4.0’ solution using piezo sensors,” *International Journal of Building Pathology and Adaptation*, vol. 39, no. 2, pp. 283–311, May 2020, doi: 10.1108/IJBPA-12-2019-0111.

[6] Y. Luo, Y. Chen, H.-P. Wan, F. Yu, and Y. Shen, “Development of laser-based displacement monitoring system and its application to large-scale spatial structures,” *J Civil Struct Health Monit*, vol. 11, no. 2, pp. 381–395, Apr. 2021, doi: 10.1007/s13349-020-00459-4.

[7] M. Abdulkarem, K. Samsudin, F. Z. Rokhani, and M. F. A Rasid, “Wireless sensor network for structural health monitoring: A contemporary review of technologies, challenges, and future direction,” *Structural Health Monitoring*, vol. 19, no. 3, pp. 693–735, May 2020, doi: 10.1177/1475921719854528.

[8] V. R. Gharehbaghi *et al.*, “A Critical Review on Structural Health Monitoring: Definitions, Methods, and Perspectives,” *Arch Computat Methods Eng*, vol. 29, no. 4, pp. 2209–2235, Jun. 2022, doi: 10.1007/s11831-021-09665-9.

[9] D. Goyal and B. S. Pabla, "The Vibration Monitoring Methods and Signal Processing Techniques for Structural Health Monitoring: A Review," *Arch Computat Methods Eng*, vol. 23, no. 4, pp. 585–594, Dec. 2016, doi: 10.1007/s11831-015-9145-0.

[10] S. Hassani and U. Dackermann, "A Systematic Review of Advanced Sensor Technologies for Non-Destructive Testing and Structural Health Monitoring," *Sensors*, vol. 23, no. 4, p. 2204, Feb. 2023, doi: 10.3390/s23042204.

[11] A. Moreno-Gomez, C. A. Perez-Ramirez, A. Dominguez-Gonzalez, M. Valtierra-Rodriguez, O. Chavez-Alegria, and J. P. Amezquita-Sanchez, "Sensors Used in Structural Health Monitoring," *Arch Computat Methods Eng*, vol. 25, no. 4, pp. 901–918, Nov. 2018, doi: 10.1007/s11831-017-9217-4.

[12] A. B. Noel, A. Abdaoui, T. Elfouly, M. H. Ahmed, A. Badawy, and M. S. Shehata, "Structural Health Monitoring Using Wireless Sensor Networks: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1403–1423, 2017, doi: 10.1109/COMST.2017.2691551.

[13] T. Khuc, T. A. Nguyen, H. Dao, and F. N. Catbas, "Swaying displacement measurement for structural monitoring using computer vision and an unmanned aerial vehicle," *Measurement*, vol. 159, p. 107769, Jul. 2020, doi: 10.1016/j.measurement.2020.107769.

[14] Y. Lee, G. Lee, D. S. Moon, and H. Yoon, "Vision-based displacement measurement using a camera mounted on a structure with stationary background targets outside the structure," *Structural Control and Health Monitoring*, vol. 29, no. 11, p. e3095, 2022, doi: 10.1002/stc.3095.

[15] T. Liu, Y. Lei, and Y. Mao, "Computer Vision-Based Structural Displacement Monitoring and Modal Identification with Subpixel Localization Refinement," *Advances in Civil Engineering*, vol. 2022, no. 1, p. 5444101, 2022, doi: 10.1155/2022/5444101.

[16] J. Won, J.-W. Park, K. Park, H. Yoon, and D.-S. Moon, "Non-Target Structural Displacement Measurement Using Reference Frame-Based Deepflow," *Sensors*, vol. 19, no. 13, p. 2992, Jan. 2019, doi: 10.3390/s19132992.

[17] Y.-S. Yang, C. Wu, T. T. C. Hsu, H.-C. Yang, H.-J. Lu, and C.-C. Chang, "Image analysis method for crack distribution and width estimation for reinforced concrete structures," *Automation in Construction*, vol. 91, pp. 120–132, Jul. 2018, doi: 10.1016/j.autcon.2018.03.012.

[18] Z. Aliansyah, M. Jiang, T. Takaki, and I. Ishii, "High-speed Vision System for Dynamic Structural Distributed Displacement Analysis," *J. Phys.: Conf. Ser.*, vol. 1075, p. 012014, Aug. 2018, doi: 10.1088/1742-6596/1075/1/012014.

[19] T. Wu *et al.*, "Accurate structural displacement monitoring by data fusion of a consumer-grade camera and accelerometers," *Engineering Structures*, vol. 262, p. 114303, Jul. 2022, doi: 10.1016/j.engstruct.2022.114303.

[20] Z. Zhou, S. Shao, G. Deng, Y. Gao, S. Wang, and X. Chu, "Vision-based modal parameter identification for bridges using a novel holographic visual sensor," *Measurement*, vol. 179, p. 109551, Jul. 2021, doi: 10.1016/j.measurement.2021.109551.

[21] J. Guo, Y. Xiang, K. Fujita, and I. Takewaki, "Vision-Based Building Seismic Displacement Measurement by Stratification of Projective Rectification Using Lines," *Sensors*, vol. 20, no. 20, p. 5775, Jan. 2020, doi: 10.3390/s20205775.

[22] A. Cataldo *et al.*, "Advanced Video-Based Processing for Low-Cost Damage Assessment of Buildings under Seismic Loading in Shaking Table Tests," *Sensors*, vol. 23, no. 11, p. 5303, Jun. 2023, doi: 10.3390/s23115303.

[23] S. Zhao, F. Kang, J. Li, and C. Ma, "Structural health monitoring and inspection of dams based on UAV photogrammetry with image 3D reconstruction," *Automation in Construction*, vol. 130, p. 103832, Oct. 2021, doi: 10.1016/j.autcon.2021.103832.

[24] D. Kang and Y. Cha, "Autonomous UAVs for Structural Health Monitoring Using Deep Learning and an Ultrasonic Beacon System with Geo-Tagging," *Computer aided Civil Eng*, vol. 33, no. 10, pp. 885–902, Oct. 2018, doi: 10.1111/mice.12375.

[25] L. Parente *et al.*, "Image-Based Monitoring of Cracks: Effectiveness Analysis of an Open-Source Machine Learning-Assisted Procedure," *Journal of Imaging*, vol. 8, no. 2, p. 22, Feb. 2022, doi: 10.3390/jimaging8020022.

[26] J.-Y. Chou and C.-M. Chang, "Image Motion Extraction of Structures Using Computer Vision Techniques: A Comparative Study," *Sensors*, vol. 21, no. 18, p. 6248, Jan. 2021, doi: 10.3390/s21186248.

[27] F. Mangini *et al.*, "Tag recognition: A new methodology for the structural monitoring of cultural heritage," *Measurement*, vol. 127, pp. 308–313, Oct. 2018, doi: 10.1016/j.measurement.2018.06.003.

[28] M. Sangirardi, V. Altomare, S. De Santis, and G. de Felice, "Detecting Damage Evolution of Masonry Structures through Computer-Vision-Based Monitoring Methods," *Buildings*, vol. 12, no. 6, p. 831, Jun. 2022, doi: 10.3390/buildings12060831.

[29] Y.-S. Yang, "Measurement of Dynamic Responses from Large Structural Tests by Analyzing Non-Synchronized Videos," *Sensors*, vol. 19, no. 16, p. 3520, Jan. 2019, doi: 10.3390/s19163520.

[30] J.-W. Park, J.-J. Lee, H.-J. Jung, and H. Myung, "Vision-based displacement measurement method for high-rise building structures using partitioning approach," *NDT & E International*, vol. 43, no. 7, pp. 642–647, Oct. 2010, doi: 10.1016/j.ndteint.2010.06.009.

[31] P. S. Harvey Jr and G. Elisha, "Vision-based vibration monitoring using existing cameras installed within a building," *Structural Control and Health Monitoring*, vol. 25, no. 11, p. e2235, 2018.

[32] A. Z. Hosseinzadeh and P. Harvey Jr, "Pixel-based operating modes from surveillance videos for structural vibration monitoring: A preliminary experimental study," *Measurement*, vol. 148, p. 106911, 2019.

[33] A. Z. Hosseinzadeh, M. Tehrani, and P. Harvey Jr, "Modal identification of building structures using vision-based measurements from multiple interior surveillance cameras," *Engineering Structures*, vol. 228, p. 111517, 2021.

[34] J. Zhou, L. Huo, C. Huang, Z. Yang, and H. Li, "Feasibility study of earthquake-induced damage assessment for structures by utilizing images from surveillance cameras," *Structural Control and Health Monitoring*, vol. 2024, no. 1, p. 4993972, 2024.

[35] D. Feng and M. Q. Feng, “Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection – A review,” *Engineering Structures*, vol. 156, pp. 105–117, Feb. 2018, doi: 10.1016/j.engstruct.2017.11.018.

[36] B. F. Spencer, V. Hoskere, and Y. Narazaki, “Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring,” *Engineering*, vol. 5, no. 2, pp. 199–222, Apr. 2019, doi: 10.1016/j.eng.2018.11.030.

[37] C.-Z. Dong and F. N. Catbas, “A review of computer vision–based structural health monitoring at local and global levels,” *Structural Health Monitoring*, vol. 20, no. 2, pp. 692–743, Mar. 2021, doi: 10.1177/1475921720935585.

[38] C. Ferraris, G. Amprimo, and G. Pettiti, “Computer Vision and Image Processing in Structural Health Monitoring: Overview of Recent Applications,” *Signals*, vol. 4, no. 3, pp. 539–574, Jul. 2023, doi: 10.3390/signals4030029.

[39] Y. Zhuang, W. Chen, T. Jin, B. Chen, H. Zhang, and W. Zhang, “A Review of Computer Vision-Based Structural Deformation Monitoring in Field Environments,” *Sensors*, vol. 22, no. 10, p. 3789, May 2022, doi: 10.3390/s22103789.

[40] K. Luo, X. Kong, J. Zhang, J. Hu, J. Li, and H. Tang, “Computer Vision-Based Bridge Inspection and Monitoring: A Review,” *Sensors*, vol. 23, no. 18, p. 7863, Sep. 2023, doi: 10.3390/s23187863.

[41] N. Kerle, F. Nex, M. Gerke, D. Duarte, and A. Vetrivel, “UAV-Based Structural Damage Mapping: A Review,” *IJGI*, vol. 9, no. 1, p. 14, Dec. 2019, doi: 10.3390/ijgi9010014.

[42] S. Poorghasem and Y. Bao, “Review of robot-based automated measurement of vibration for civil engineering structures,” *Measurement*, vol. 207, p. 112382, Feb. 2023, doi: 10.1016/j.measurement.2022.112382.

[43] Y.-S. Yang *et al.*, “Image analysis applications for building inter-story drift monitoring,” *Applied Sciences*, vol. 10, no. 20, p. 7304, 2020.

[44] T.-Y. Hsu and X.-J. Kuo, “PDP method to compesate for rotational effect when using a single surveillance camera for interstory drift measurement,” *Measurement Science and Technology*, vol. 31, no. 9, p. 095902, 2020.

[45] W. Wen, C. Zhang, C. Zhai, J. Guo, and J. Hu, “A method for automatic monitoring structural earthquake response using surveillance video,” *Journal of Building Engineering*, p. 113737, 2025.

[46] C. Zhang, W. Wen, C. Zhai, and J. Guo, “Method of extracting structural seismic response from blurred surveillance videos based on line matching technique,” *Measurement*, vol. 248, p. 116937, 2025.

[47] W. Wen, C. Zhang, J. Hu, J. Guo, C. Zhai, and B. Zhou, “Automatic monitoring method for seismic response of building structures and equipment based on indoor surveillance cameras,” *Mechanical Systems and Signal Processing*, vol. 224, p. 112220, 2025.