Available online at www.easd.procedia.org
EASD Procedia EURODYN (2026) 1437-1446

EASD Procedia

EURODYN 2026
XIII International Conference on Structural Dynamics
M. Beer, C. Papadimitriou, H. Jensen, J. Chen, M. Broggi, J. Behrensdorf (eds.)
Hannover, Germany, 27.09.–01.10 2026

NUMERICAL STUDY ON DATA FUSION FOR STRUCTURAL MONITORING USING SPARSE ACCELEROMETERS AND VISION SYSTEMS

Ghita Eslami Varzaneh[1], Federico Ponsi[1], Elisa Bassoli[1], and Loris Vincenzi[1]

[1]University of Modena and Reggio Emilia, Department of Engineering "Enzo Ferrari"
via Vivarelli 10, 41125 Modena, Italy
e-mail: {ghita.eslami,federico.ponsi,elisa.bassoli,loris.vincenzi}@unimore.it

Keywords: Structural health monitoring, Data fusion, Multi-rate Kalman filter, Virtual sensing, Vision-based system.

Abstract. *Structural health monitoring demands vibration data with high resolution in both time and space to detect and track the dynamic behavior of the system. While accelerometers are well-established and provide accurate temporal data, their use is limited by high installation costs, the need for direct access to the structure, and integration errors in displacement estimation. Conversely, vision-based systems offer broad spatial coverage and low intrusiveness, yet suffer from limited temporal resolution, environmental sensitivity, and delicate post-processing issues. To overcome these limitations, this work introduces a data fusion framework that integrates sparse accelerometer inputs with vision-based displacement measurements. By combining these complementary sources, the method enables accurate and reliable monitoring of structural dynamics while minimizing the need for dense sensor networks. Indeed, only a few accelerometers are needed, as acceleration data at un-instrumented but visually-monitored locations is reconstructed via a virtual sensing technique based on modal superposition.*

ISSN:2311-9020 © 2026 The Authors. Published by EASD Procedia.
Peer-review under responsibility of the Organizing Committee of EURODYN 2026.
doi: 10.47964/1326.10754.074

1 INTRODUCTION

Dynamic monitoring involves assessing structural health by measuring motion-related quantities (such as displacements, velocities, or accelerations) to identify modal parameters, whose changes can serve as potential indicators of structural damage. Traditionally, accelerometer systems have been widely adopted due to their well-established accuracy and robustness under various conditions [1]. Despite providing high temporal resolution, their spatial resolution is limited by high installation costs, which increase with the number of measurement points, and by the requirement for direct access to the structure and extensive wiring. Furthermore, displacement estimation requires double integration of acceleration measurements, which can reduce the accuracy of long-term or low-frequency displacement estimates [2]. As a result, neither permanent displacements resulting from structural damage nor pseudo-static displacements induced by moving loads can be reliably detected using accelerometers.

On the other hand, vision-based systems are attracting increasing attention as non-contact, cost-effective solutions capable of monitoring multiple points simultaneously with a single consumer-grade camera [3]. They also provide direct displacement measurements, avoiding the accumulation errors associated with double integration of acceleration data. However, these systems are sensitive to environmental factors (e.g., lighting conditions, weather, occlusions, air refraction) as well as to unintended camera movements (caused by wind or user handling), which might be misinterpreted as motion of the structure itself. Furthermore, accurately converting pixel measurements into real-world units remains a delicate task that directly affects data quality when a single camera is employed [4]. Lastly, detecting low-amplitude or high-frequency vibrations remains challenging due to limitations in camera resolution and frame rate.

The integration of complementary sensing techniques is commonly referred to as data fusion [5]. In feature-level fusion, heterogeneous physical measurements (such as accelerations and displacements) are combined [6]. This can involve integrating acceleration data from conventional sensors with displacement measurements obtained from vision-based systems, although applications of accelerometer-vision data fusion remain relatively limited in the literature [7–9]. In the present work, this approach is implemented using a multi-rate Kalman filter [10, 11] to reconcile differing sampling frequencies, as typically occurs with accelerometer- and vision-based data. This enables the capture of both low- and high-frequency vibrations. By combining the high temporal resolution of accelerometers with the high spatial resolution of vision-based measurements, their complementary strengths improve the reliability and interpretability of structural monitoring.

Only a limited number of accelerometers are employed in the setup to ensure minimal installation effort, reduced invasiveness, and lower costs. To compensate for the sparse instrumentation, virtual sensing is adopted to estimate acceleration responses at locations where direct measurements are not available. In particular, a modal-based virtual sensing approach is used to accurately predict accelerations at un-instrumented points through modal expansion. This strategy provides a cost-effective and reliable solution for dynamic monitoring, bridging the gap between densely instrumented sensor networks and fully vision-based systems.

The proposed framework is validated through numerical simulations on a footbridge. Displacement and acceleration responses are generated along the structure under multi-harmonic loading, here represented as a pedestrian traversing the example footbridge. Unlike white-noise excitation, the spatially moving pedestrian induces low-frequency displacement drifts in the structural response, information that would be lost if only accelerometers were used. A subset of the simulated acceleration data is treated as unavailable and subsequently reconstructed

through virtual sensing. Both displacement measurements and reconstructed accelerations are then subjected to increasing levels of noise and combined through data fusion. Finally, the fused results are compared with the original noise-free simulated data to assess the accuracy of the method as a function of measurement uncertainty. These analyses numerically confirm the enhancement obtained through the fusion of vision-based displacement measurements and accelerometer data.

The paper is organised as follows. The multi-rate Kalman filter-based data fusion is introduced in Section 2. The modal-based virtual sensing technique is presented in Section 3. Numerical analyses are conducted in Section 4. Conclusions are drawn in Section 5.

2 DATA FUSION

This Section presents a Kalman filter for data fusion [10, 11], suited for combining displacement and acceleration measurements acquired at different sampling rates. The system state vector $\boldsymbol{x} = [x\ \dot{x}]^\top$ includes displacement x and velocity $\dot{x}$, while the input $\ddot{u}$ corresponds to measured acceleration and the output z to measured displacement (i.e. vision-based data). For practical implementation with acceleration sampled at T_a, the system is discretized at discrete time instants $t_k = kT_a$:

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{A}_d \boldsymbol{x}_k + \boldsymbol{B}_d \ddot{u}_k + \boldsymbol{w}_k \\ z_k &= \boldsymbol{H}\boldsymbol{x}_k + v_k \end{aligned} \tag{1}$$

The system matrices $\boldsymbol{A}_d$ and $\boldsymbol{B}_d$ describe the discrete-time evolution of the states driven by acceleration, whereas $\boldsymbol{H}$ selects the measured component (displacement). The process noise $\boldsymbol{w}_k$, with covariance $\boldsymbol{Q}_d$, accounts for uncertainty in the acceleration input and is therefore non-zero only in the acceleration-related state, while the measurement noise v_k, with covariance $\boldsymbol{R}_d$, represents uncertainty in displacement measurements. The Kalman filter is initialized with initial state estimate $\hat{\boldsymbol{x}}_0$ and initial error covariance $\boldsymbol{P}_0$, usually set with zero mean and identity covariance matrix. Then, it recursively estimates the state $\hat{\boldsymbol{x}}$ and its covariance $\boldsymbol{P}$ balancing model predictions and available measurements, through prediction (time update) and correction (measurement update) steps. In multi-rate conditions, where accelerations are sampled more frequently than displacements, only the prediction step is performed between two displacement measurements, while the correction is applied when new displacement data become available. In multi-rate scenarios, where accelerations are sampled more frequently than displacements ($T_d > T_a$), this approach reconciles differing sampling intervals. Post-processing smoothing is applied to reduce noise and improve temporal consistency. For a detailed treatment, please see [10, 11].

3 VIRTUAL SENSING

This Section introduces a modal-based virtual sensing approach that enables operation with a limited number of accelerometers and the reconstruction of acceleration responses at uninstrumented, yet structurally relevant, locations where vision-based displacement measurements are available for data fusion. The approach is based on modal superposition, combining experimentally identified modal parameters with analytically approximated mode shapes.

First, the acceleration signals $\ddot{u}_{\text{ref}}$ measured at the N sensor locations $\boldsymbol{l}_{\mathbf{ref}} = [l_{\text{ref},1}, \ldots, l_{\text{ref},N}]$ are processed through the Enhanced Frequency Domain Decomposition (EFDD) [12] method to identify modal properties. For each mode n, the corresponding natural frequency f_n, damping ratio ξ_n, and mode shape components $\phi_n(\boldsymbol{l}_{\mathbf{ref}})$ at the instrumented points are extracted. The

Modal Assurance Criterion (MAC) is employed to validate the consistency of mode shapes across frequency bands, ensuring the reliability of the identified modes.

To separate the contribution of each mode, Band-Pass Filters (BPF) are applied to the signal recorded at position i (with $i = 1, \ldots, N$), $\ddot{u}_{\text{ref}}(l_{\text{ref},i})$. The filters are centred on the identified natural frequencies and restricted to frequency ranges with acceptable MAC values, yielding the modal acceleration component at the i-th sensor for mode n:

$$\ddot{u}_{\text{ref},n}(l_{\text{ref},i}) = \text{BPF}[\ddot{u}_{\text{ref}}(l_{\text{ref},i}), f_n] \tag{2}$$

Assuming the footbridge as a simply supported beam with a planar rather than 3-dimensional structural behaviour, the mode shapes $\phi_n(l)$ are approximated by sine functions along the structural coordinate l over the beam length L:

$$\phi_n(l) = \alpha_n \sin(\beta_n l + \gamma_n) \tag{3}$$

Based on the known mode shape components at $\boldsymbol{l}_{\mathbf{ref}}$, the amplitude α_n, the spatial frequency term β_n and the phase shift γ_n are estimated through least-squares curve fitting. Although fitting three parameters would typically require three physical sensors (i.e., three known mode shape components), two sensors are sufficient for the estimation under the assumption of a simply supported beam with zero mode shapes at the supports.

The modal acceleration at an un-instrumented location l_{est} is estimated by scaling the filtered acceleration at a reference sensor $l_{\text{ref},i}$, according to the relative mode shape amplitudes:

$$\ddot{u}_{\text{est},n}(l_{\text{est}}) = \ddot{u}_{\text{ref},n}(l_{\text{ref},i}) \frac{\phi_n(l_{\text{est}})}{\phi_n(l_{\text{ref},i})} \tag{4}$$

The overall acceleration $\ddot{u}_{\text{est}}$ at location l_{est} is finally reconstructed through modal superposition.

4 NUMERICAL ANALYSES

This Section validates the proposed approach through numerical simulations. The example footbridge is a simply supported 50-m beam, discretized in 1-m increments along its length. Four sinusoidal modes are considered, with damping ratios of 0.5% and natural frequencies chosen as 1, 2, 3, 4 Hz (Case #1) and 2, 3, 4, 5 Hz (Case #2). The total structural mass is 25000 kg, assumed uniformly distributed across the discretized points to assemble the mass matrix, from which the modal masses are computed using the corresponding mode shapes.

As a representative sinusoidal multi-harmonic excitation, a pedestrian-induced loading is modelled as:

$$P(t) = G + \sum_{m=1}^{4} G\,\text{DLF}_m \sin(2\pi m f_s t) \tag{5}$$

expressed as the superposition of four harmonic components corresponding to the pacing frequency f_s and its integer multiples. The dynamic load factors DLF_m are f_s-dependent coefficients defined according to [13]. A typical walking speed under undisturbed conditions is selected ($v_s = 1.34$ m/s as reported in [14]), corresponding to a pacing frequency of $f_s = 1.91$ Hz consistent with [15]. The pedestrian load is defined by a body weight G, computed from a mass of 73.85 kg (average value reported in [16]), and is assumed to travel along the structure at a constant velocity v_s.

For each mode, the modal force is obtained by projecting the load onto the corresponding mode shape, and the resulting modal accelerations and displacements are evaluated at selected locations along the structure ($L/2$, $L/3$, $L/4$, $L/5$, $L/6$). The total response at a specific position is then reconstructed via modal superposition, with accelerations sampled at 80 Hz and displacements resampled at 30 Hz (i.e., 30 frames per second), to reflect typical sampling rates of vision-based systems.

To emulate a setup with sparse instrumentation, only the acceleration measurements at $L/6$ and $L/5$ are assumed to be available and are used as references to reconstruct the response at $L/4$ via modal-based virtual sensing (see Section 3). To assess the impact of measurement uncertainty, the analysis is repeated by contaminating the reference signals with Gaussian white noise, whose amplitude is defined as a percentage of the maximum absolute value of the signal (ranging from 5% to 25% in 5% increments). With increasing noise in the reference signals $\ddot{u}_{\text{ref}}$, the reconstructed acceleration $\ddot{u}_{\text{est}}$ at $L/4$ is compared against the corresponding noise-free simulated response $\ddot{u}$. For Case #1, Figure 1(a) and Figure 1(b) show the extreme scenarios with 0% and 25% noise in the time domain, respectively, while Figure 2(a) and Figure 2(b) present the corresponding scenarios in the frequency domain. These results demonstrate that the approach maintains good reconstruction performance even when the reference signals are affected by noise, as commonly occurs in real-world conditions.

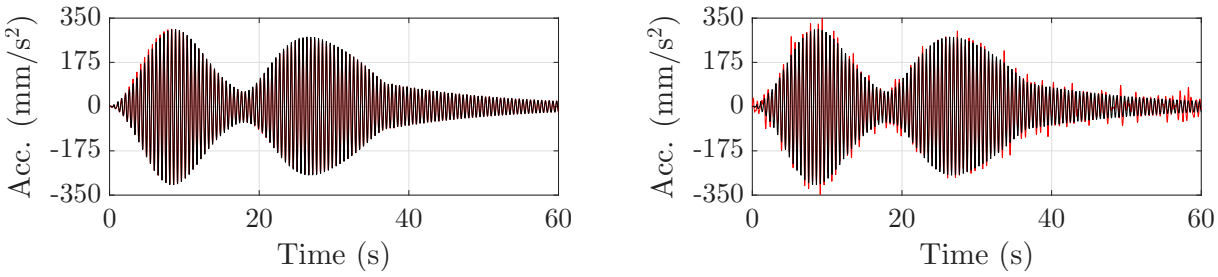


Figure 1: Case #1: Time-domain comparison of simulated accelerations at $L/4$, in black, with reconstructed accelerations, in red, for (a) noise-free condition and (b) 25% noise.

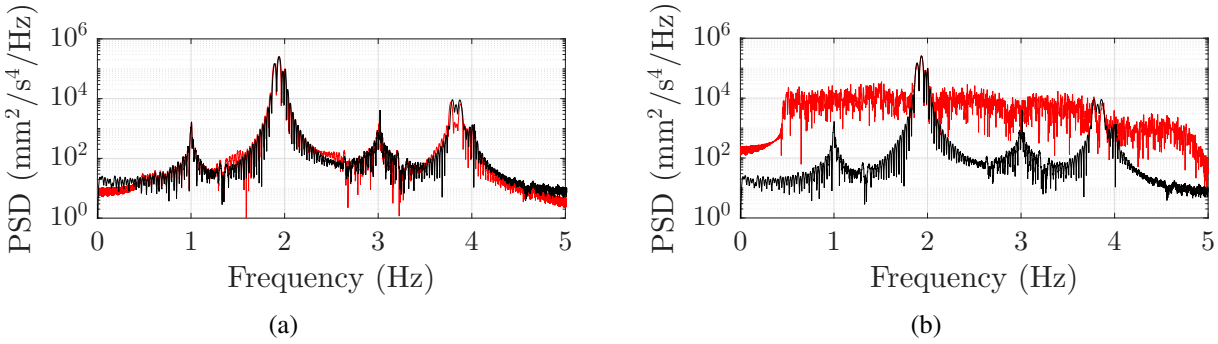


Figure 2: Case #1: Frequency-domain comparison (in logarithmic scale) of simulated accelerations at $L/4$, in black, with reconstructed accelerations, in red, for (a) noise-free condition and (b) 25% noise.

	Noise level					
Location	0%	5%	10%	15%	20%	25%
$L/4$	0.94	1.79	3.31	4.68	6.54	8.07
$L/3$	2.18	2.87	4.57	6.22	8.44	10.26
$L/2$	21.18	25.04	34.11	45.52	55.34	72.70

Table 1: Case #1: NRMSE (%) between simulated and reconstructed accelerations for different noise levels and prediction locations.

	Noise level					
Location	0%	5%	10%	15%	20%	25%
$L/4$	1.21	1.99	3.42	4.94	6.28	8.14
$L/3$	2.13	2.63	3.75	5.26	6.46	8.31
$L/2$	2.83	3.08	3.62	4.38	5.36	5.94

Table 2: Case #2: NRMSE (%) between simulated and reconstructed accelerations for different noise levels and prediction locations.

As concerns Case #1, the Root Mean Square Error normalized to the signal peak amplitude (NRMSE) between the simulated (noise-free) and reconstructed time series is reported in Table 1 for all the examined noise levels. The NRMSE values indicate that prediction errors increase with noise but remain limited, demonstrating the effectiveness of virtual sensing. The same procedure is repeated to predict the accelerations at $L/3$ and $L/2$, with the corresponding NRMSE values also included in Table 1. Since the reconstruction relies on the superposition of a limited number of modes, some modal contributions may vanish at specific locations (i.e., nodal points), which can become critical when the corresponding modes are significantly excited by the forcing. Indeed, when the excitation activates the second mode (whose modal contribution is zero at midspan), as in Case #1, the errors at $L/2$ become significantly large.

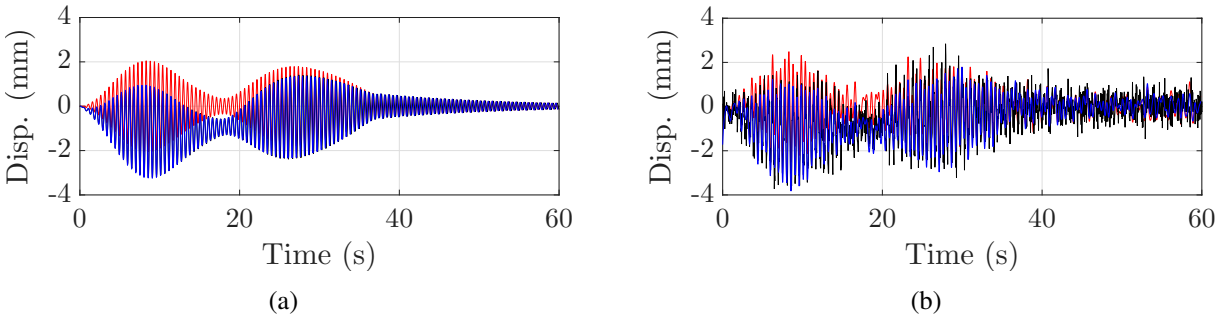


Figure 3: Case #1: Time-domain comparison displacements at $L/4$ from Kalman filter-based data fusion, in blue, displacements obtained by double integration of reconstructed accelerations, in red, and simulated displacements, in black, for (a) noise-free condition and (b) 25% noise.

Table 2 reports the NRMSE values for Case #2. Unlike Case #1, this scenario assumes that the pacing frequency is approximately aligned with the first vibration mode rather than the second.

		Noise level					
Location	Type	0%	5%	10%	15%	20%	25%
$L/4$	DF-acc	14.83	14.86	15.39	15.58	16.92	17.67
	DF-disp	1.30	2.10	3.87	5.48	7.88	9.58
$L/3$	DF-acc	18.84	18.95	19.04	19.78	21.02	20.45
	DF-disp	1.70	2.57	4.39	6.05	9.03	10.23
$L/2$	DF-acc	36.55	37.77	40.19	43.13	45.49	50.45
	DF-disp	4.25	10.46	17.92	23.16	28.09	35.11

Table 3: Case #1: NRMSE (%) between data fusion estimates and displacement obtained by double integration of noise-free simulated accelerations (DF-acc) and noise-free simulated displacements (DF-disp), for different noise levels and locations.

		Noise level					
Location	Type	0%	5%	10%	15%	20%	25%
$L/4$	DF-acc	9.77	9.99	10.37	10.93	11.46	12.95
	DF-disp	1.80	2.44	3.74	5.29	6.57	8.22
$L/3$	DF-acc	8.76	8.97	9.34	9.99	10.35	11.31
	DF-disp	2.15	2.67	3.87	4.88	6.74	8.38
$L/2$	DF-acc	8.13	8.39	8.69	9.37	10.34	11.06
	DF-disp	3.01	3.36	4.30	5.36	6.95	8.44

Table 4: Case #2: NRMSE (%) between data fusion estimates and displacement obtained by double integration of noise-free simulated accelerations (DF-acc) and noise-free simulated displacements (DF-disp), for different noise levels and locations.

Consequently, the errors at $L/2$ remain limited and are of the same order of magnitude as those at other locations, namely $L/4$ and $L/3$. This is a key aspect to be considered in the design of the experimental setup, namely the selection of un-instrumented locations for which the acceleration signals are to be reconstructed, based on the expected excitation frequency. When this aspect is properly taken into account (e.g., Case #2), the reconstruction procedure achieves significant accuracy even as data noise increases, with errors that remain bounded compared to the noise level assigned to the input signals, at all locations along the structure.

Afterwards, simulated displacements are calculated and subjected to the same levels of Gaussian white noise as the acceleration signals. Specifically, the noise-free simulated displacements represent the exact displacements, whereas noisy displacements represent pseudo-experimental data, here corresponding to vision-based measurements (i.e., z in Equation 2). The reconstructed accelerations and the noisy displacements are then combined through data fusion as indicated in Section 2. Figure 3 illustrates the example case at $L/4$ (Case #1), comparing three displacement time series: that obtained by double integration of the reconstructed acceleration, u_{est}, the simulated noisy displacement (here representing vision-based data), z, and the estimate derived from Kalman-filter-based data fusion, x. Results are shown for both noise-free condition and 25% noise (Figure 3(a) and Figure 3(b), respectively), highlighting the effectiveness of combining vision-based displacements with accelerometer measurements as complementary sources of information.

Indeed, modal-based virtual sensing leverages the spatial resolution of vision-based measurements without requiring dense sensor networks. Furthermore, the data fusion approach exploits the high temporal resolution of accelerations and the complementary sensitivity of the two measured quantities (i.e., displacements and accelerations) to low- and high-frequency components of the structural response. To provide quantitative support, NRMSE values are reported in Table 3 and Table 4 for Case #1 and Case #2, respectively, including all noise levels and all investigated locations. Specifically, the NRMSE is computed between the data fusion results (for each noise level) and the noise-free simulated displacements (DF-disp), as well as between the data fusion results and the displacements u obtained by double integration of noise-free simulated accelerations $\ddot{u}$ (DF-acc).

Table 3 and Table 4 show that using acceleration alone fails to capture low-frequency drifts (i.e., resulting in high DF-acc NRMSE values), particularly when the configuration of the limited number of sensors is not properly designed with respect to the excited modes, as in Case #1. This clearly highlights the effectiveness of accelerometer-vision data fusion for achieving accurate dynamic monitoring, with the results demonstrating that data fusion acts as a denoising procedure. Overall, the proposed procedure leads to an error reduction of over 50%.

5 CONCLUSIONS

This study presented a data fusion framework for structural dynamic monitoring, combining vision-based displacement measurements with accelerometer data through a Kalman filter. In addition, a modal-based virtual sensing approach was adopted to reconstruct acceleration signals from a limited number of measurements, enabling the estimation of responses at un-instrumented locations.

The results demonstrated that the reconstruction approach provides accurate estimation of acceleration time series. Moreover, data fusion effectively combines low- and high-frequency information from different sources. By integrating heterogeneous sensing techniques, the method leverages their complementary strengths, including the high temporal resolution of accelerometers and the spatial richness of vision-based measurements.

A systematic assessment under increasing levels of noise showed that the framework remains robust under realistic noisy conditions. While prediction errors increase with noise, they remain bounded, confirming the stability of the method. More specifically, when the configuration of the limited number of sensors is properly designed with respect to the excited modes, the method enables a reduction of noise by more than 50% compared to measured signals.

Overall, the proposed framework offers a cost-effective and flexible solution for structural monitoring, reducing the need for dense sensor networks while maintaining reliable performance. Future work will focus on experimental validation under real-world conditions and on extending the approach to more complex scenarios.

REFERENCES

[1] Hassani Sahar and Dackermann Ulrike. “A systematic review of advanced sensor technologies for non-destructive testing and structural health monitoring”. In: *Sensors* 23.4 (2023), p. 2204. DOI: 10.3390/s23042204.

[2] Gavin Henri P., Morales Rodrigo, and Reilly Kathryn. “Drift-free integrators”. In: *Review of Scientific Instruments* 69.5 (1998), pp. 2171–2175. DOI: 10.1063/1.1148918.

[3] Zona Alessandro. “Vision-based vibration monitoring of structures and infrastructures: An overview of recent applications”. In: *Infrastructures* 6.1 (2020), p. 4. DOI: 10.3390/infrastructures6010004.

[4] Ponsi Federico et al. “An automated vision-based approach for structural dynamic identification”. In: *Journal of Civil Structural Health Monitoring* 16.2 (2026), p. 24. DOI: 10.1007/s13349-025-01062-1.

[5] Wu Rih-Teng and Jahanshahi Mohammad Reza. “Data fusion approaches for structural health monitoring and system identification: Past, present, and future”. In: *Structural Health Monitoring* 19.2 (2020), pp. 552–586. DOI: 10.1177/1475921718798769.

[6] Ponsi Federico et al. “Structural displacement estimation through data fusion of GNSS and acceleration data”. In: *Operational Modal Analysis*. CRC Press, pp. 26–41.

[7] Park Jong-Woong et al. “Visual–inertial displacement sensing using data fusion of vision-based displacement with acceleration”. In: *Structural Control and Health Monitoring* 25.3 (2018), e2122. DOI: 10.1002/stc.2122.

[8] Ma Zhanxiong et al. “Structural displacement estimation by fusing vision camera and accelerometer using hybrid computer vision algorithm and adaptive multi-rate kalman filter”. In: *Automation in Construction* 140 (2022), p. 104338. DOI: 10.1016/j.autcon.2022.104338.

[9] Ponsi Federico et al. “Combining displacement and acceleration data for structural health assessment”. In: *Proceedings of the 19th ECCOMAS Thematic Conference on Computational Methods in Structural Dynamics and Earthquake Engineering*. Rhodes Island, Greece, 2025, pp. 4913–4922. DOI: 10.7712/120125.12783.26002.

[10] Smyth Andrew and Meiliang Wu. “Multi-rate Kalman filtering for the data fusion of displacement and acceleration response measurements in dynamic system monitoring”. In: *Mechanical Systems and Signal Processing* 21.2 (2007), pp. 706–723. DOI: 10.1016/j.ymssp.2006.03.005.

[11] Kim Junhee, Kiyoung Kim, and Hoon Sohn. “Autonomous dynamic displacement estimation from data fusion of acceleration and intermittent displacement measurements”. In: *Mechanical Systems and Signal Processing* 42.1 (2014), pp. 194–205. DOI: 10.1016/j.ymssp.2013.09.014.

[12] Brincker Rune, Ventura Carlos E., and Andersen Palle. “Damping estimation by frequency domain decomposition”. In: *Proceedings of the 19th International Modal Analysis Conference*. Orlando, FL, USA, 2001, pp. 698–703.

[13] Young P. “Improved floor vibration prediction methodologies”. In: *ARUP Vibration Seminar*. 2001.

[14] Buchmüller Stefan and Weidmann Ulrich. “Parameters of pedestrians, pedestrian traffic and walking facilities”. In: *IVT Schriftenreihe* 132 (2006). DOI: 10.3929/ethz-b-000047950.

[15] Bruno Luca and Venuti Fiammetta. “The pedestrian speed-density relation: Modelling and application”. In: *Proceedings of the 3rd International Conference on Footbridges*. Porto, Portugal: FEUP, 2008, pp. 255–256.

[16] Portier Kenneth, Keith Tolson J., and Roberts Stephen M. “Body weight distributions for risk assessment”. In: *Risk Analysis* 27.1 (2007), pp. 11–26. DOI: 10.1111/j.1539-6924.2006.00856.x.