Journal of Civil Structural Health Monitoring (2026) 16:108
https://doi.org/10.1007/s13349-026-01153-7

ORIGINAL PAPER

Rapid phase-based displacement measurement based on complex orthogonal wavelet transform for vision-based structural response monitoring

Sida Ai[1] · **Chuanzhi Dong**[2] · **Qipei Mei**[1]

Received: 30 March 2026 / Revised: 24 September 2026 / Accepted: 8 October 2026
© Springer-Verlag GmbH Germany, part of Springer Nature 2026

Abstract

Displacement is particularly crucial in civil infrastructure monitoring, where precise measurement of structural movements helps assess stability and detect potential failures. In the field of vision-based displacement measurement, phase-based optical flow is receiving increasing attention for its robustness under illumination variation and high subpixel accuracy. However, the high computational cost is always viewed as a significant drawback of this method. This issue primarily arises from the overcompleteness of the computation process in existing phase extraction methods, which are crucial in phase-based displacement measurement. To address this problem, a new phase extraction method, termed the complex orthogonal wavelet transform (COWT), is proposed using the non-redundant decomposition framework of the orthogonal wavelet transform to improve computational efficiency. The proposed method is evaluated through a laboratory shaker test on a column and a full-scale outdoor shake table test on a cold-formed steel wall structure, yielding mean absolute errors (MAEs) of 0.3147 mm and 1.8945 mm, respectively. Although these errors are higher than those obtained using traditional phase extraction methods, COWT achieves MAEs that are 13.4% and 60.6% lower than those of the best-performing recently developed phase extraction method in each experiment, respectively. Importantly, COWT is at least 10 times faster than traditional phase extraction methods and substantially faster than other recently developed phase extraction methods, demonstrating a favorable balance between measurement accuracy and computational efficiency.

Keywords Phase-based optical flow · New phase extraction method · Complex orthogonal wavelet transform · Non-redundant decomposition · Computational efficiency

1 Introduction

During operation, civil structures are subjected to various external actions, such as traffic, earthquakes, and wind. Tracking their structural behavior in both experimental and real-world environments is essential for assessing their condition and ensuring safe operation [1]. Conventional displacement measurement approaches can generally be grouped into contact-based and non-contact approaches. Representative contact-based sensors include linear variable differential transformers (LVDTs) and pull-wire displacement sensors [2, 3]. Their primary limitations include the demanding installation process and the risk of altering the structure through extra mass or externally applied forces. By comparison, non-contact approaches do not suffer from these shortcomings. GPS-based techniques are a representative example, although their precision is generally confined to the centimeter range [4]. Laser displacement sensors can achieve high accuracy, but they are relatively expensive [5]. On the other hand, vision-based methods provide a cost-effective measurement solution and can deliver accurate and reliable results when integrated with suitable computer vision algorithms [6, 7].

Digital image correlation (DIC) is a widely used displacement measurement algorithm in vision-based methods, offering full-field displacement and strain measurements at

✉ Qipei Mei
qipei.mei@ualberta.ca

[1] Department of Civil and Environmental Engineering, University of Alberta, Edmonton, AB T6G 2W2, Canada

[2] Department of Computer Science, The University of Texas at Austin, Austin, TX 78712, USA

Published online: 26 October 2026

pixel-level accuracy [8]. However, its performance is sensitive to image quality and requires labor-intensive surface patterning [9]. Compared to DIC, optical flow methods, which compute the velocity of image pixels across consecutive frames [10], provide an alternative by estimating displacement based on pixel intensity changes, often achieving subpixel accuracy without the need for pre-applied patterns. To address the aperture problem, which arises when motion is observed through a small window and leads to ambiguity in determining the true motion direction, the Lucas-Kanade (LK) method was introduced, assuming uniform displacement within small regions [11]. Optical flow is also frequently employed in vision-based structural monitoring. For example, Javh et al. [12] proposed a simplified version to achieve high-accuracy subpixel displacement measurements and validated it experimentally. Bhowmick et al. [13] used a Lagrangian representation of continuous edges in beam and frame structures for displacement tracking. Lydon et al. [14, 15] combined LK optical flow with feature extraction to measure displacements of various bridges, such as truss, pre-stressed concrete, and cable-stayed steel footbridges. A recent development in optical flow is the deep-learning-based Recurrent All-Pairs Field Transforms (RAFT), which builds a dense correlation volume to encode matching costs for every pixel pair across two consecutive frames [16]. By iteratively refining the flow through a recurrent update module, RAFT achieves high accuracy and performs well even in the presence of large motions or complex textures. Although optical flow can achieve high subpixel accuracy, it is sensitive to noise, computationally intensive, and relies on brightness constancy. These factors can reduce accuracy in dynamic or unevenly lit conditions.

In recent years, vision-based techniques for displacement measurement have continued to evolve, with increasing adoption of phase-based methods in structural health monitoring. These approaches mainly rely on phase-based optical flow, in which displacement is extracted from the phase components of image frames. Compared to intensity-based optical flow, phase-based optical flow is demonstrated to offer high subpixel accuracy and enhanced robustness against variations in illumination [17–20].

The phase-based optical flow was first proposed by Fleet et al. [17]. They employed a Gabor wavelet to convolve with the image, producing phase information, and utilized the optical flow equation in the phase direction to compute the component velocity of pixels along the direction of the Gabor wavelet. In 2013, Wadhwa et al. [21] employed the complex steerable pyramid (CSP) method to generate the phase information for the phase-based motion magnification method. Since then, Chen et al. [22, 23] applied the CSP approach together with the concept of phase-based optical flow to structural displacement measurement in both laboratory tests and on-site experiments, demonstrating strong measurement accuracy. Sarrafi et al. [24] adopted Gabor wavelet for motion magnification and modal identification of wind turbine blades in vision-based experimental studies. In addition, the phase-based method was shown to achieve displacement measurement with high subpixel-level precision [19, 25]. Miao et al. [20] adopted the Gabor wavelet method for the phase-based displacement measurement of a frame structure in a laboratory test, showing great subpixel measurement accuracy and robustness under illumination changes. The CSP method has been employed in the field of vision-based structural health monitoring, specifically for identifying modal frequencies and modal shapes [26]. The Hilbert transform was introduced for phase-based displacement measurement, together with the Riesz transform for two-dimensional motion analysis [27, 28]. In addition, a derivative-enhanced phase-based method was proposed for fast and accurate displacement measurement [29].

Despite its clear advantages, phase-based displacement measurement still suffers from limited computational efficiency [30–32]. This issue is particularly important in vision-based monitoring, where algorithmic efficiency affects applications including real-time sensing, high-resolution video processing, and long-term displacement measurement [7].

Generally, the phase-based measurement procedure includes two major parts: phase extraction and phase-based optical flow computation [17]. Of these two parts, phase extraction consumes the largest amount of computational resources because filtering operations must be performed on every frame [17, 33]. In practice, different algorithms may be adopted for phase extraction depending on the measurement strategy [17, 23].

Commonly used phase extraction methods for displacement measurement include the Gabor wavelet and the CSP [19, 20, 22, 24]. The Gabor wavelet is a typical continuous wavelet transform (CWT) method, featuring a redundant representation of the original signal [34]. In the case of the CSP method, its development was motivated by the goal of achieving translation- and rotation-invariant image decomposition [33]. Although the framework of the CSP method is quite similar to that of the orthogonal wavelet transform (OWT) method, as both involve high-pass and low-pass filtering as well as downsampling steps, CSP remains an overcomplete transform method because of the convolution between subband signals and steerable filters at each scale level [32]. This redundant representation or overcomplete transform process is the primary cause of computational inefficiency in these phase extraction methods [29].

Besides these two widely used phase extraction methods, researchers have recently proposed new phase extraction

methods to address the computational inefficiency of traditional phase extraction methods. Liu et al. [27] introduced the Hilbert-transform-enhanced phase-based displacement measurement method to avoid the complex parameter selection in the Gabor wavelet method. Although the simplified process for the phase extraction method may improve the computational efficiency compared to traditional methods, this method still involves bandpass filtering for all frames of the video, which also leads to a high computational burden. Li et al. [28] proposed the Riesz-transform-enhanced phase-based displacement measurement method to estimate 2D motion and improve computational efficiency. This method achieves significantly higher computational efficiency than the Gabor wavelet method, which is constrained by unidirectional computation. However, it may still show only limited improvement in computational efficiency compared to CSP, which is also suitable for multi-directional displacement measurement, because of its bandpass filtering process

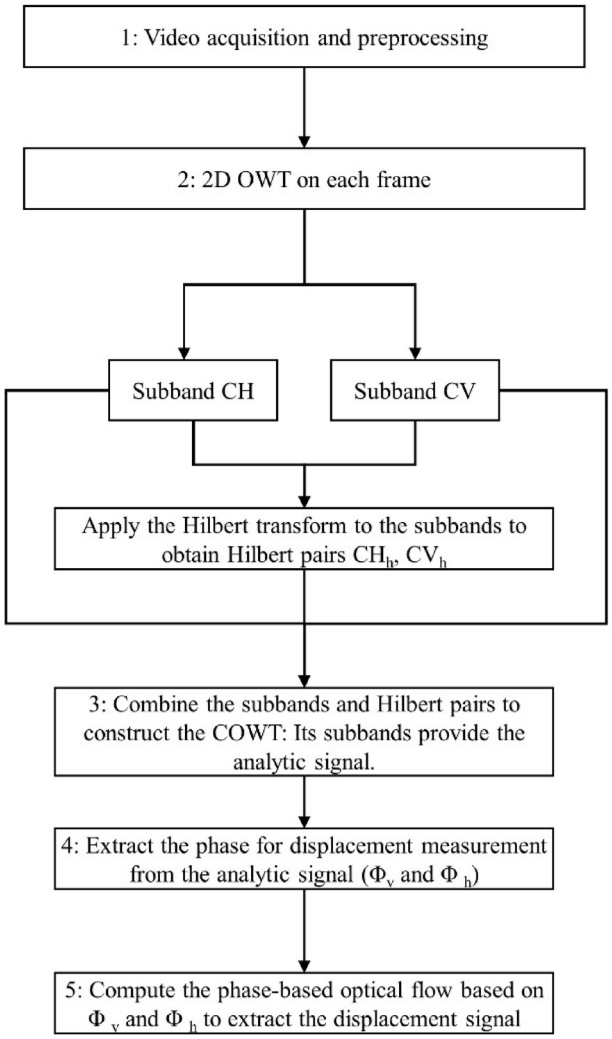


Fig. 1 The process of the COWT method for phase-based displacement measurement

applied to all frames. To further improve computational efficiency in phase-based methods, Liu et al. [29] proposed the derivative-enhanced phase-based optical flow method for fast and accurate displacement measurement. This approach generates analytical signals for phase computation by treating each raw grayscale image as the real part and its first-order derivative as the imaginary part. While this method significantly improves computational efficiency, the over-simplified phase extraction process may reduce its robustness across various scenarios.

To significantly reduce the computational cost of the phase-based method while maintaining accuracy and robustness, a practical and efficient phase-based displacement measurement method, namely the complex orthogonal wavelet transform (COWT) based on OWT, is proposed. This method is built on a non-redundant OWT decomposition framework, which significantly speeds up the phase extraction process for phase-based displacement measurement. It is at least 10 times faster than traditional phase extraction methods and also significantly faster than recently developed phase extraction methods, showing strong potential for future real-time displacement measurement.

The structure of this paper is outlined as follows: Sect. 2 presents the methodology of the proposed approach, including the decomposition process of the OWT in its two-dimensional form within a multiresolution context. Subsequently, the construction of the COWT, which primarily employs the Hilbert transform, is described. Finally, the displacement extraction procedure based on the phase information obtained from the COWT is introduced. Section 3 employs laboratory and outdoor tests to assess the effectiveness and computational efficiency of the COWT phase-based optical flow method.

2 Methodology

To overcome the efficiency and robustness limitations of the existing phase extraction methods discussed in Sect. 1, a Haar-based OWT is adopted as the underlying decomposition framework of the proposed phase extraction method, COWT. The OWT provides a non-redundant decomposition of the original signal, and the COWT is constructed by taking each OWT subband as the real part and its Hilbert transform pair as the imaginary part, thereby supplying the analytic signal required for phase extraction. This section introduces the decomposition process of the 2D OWT, the construction of the COWT and the phase extraction, and the displacement computation process.

Figure 1 presents the overall workflow of the proposed COWT approach for phase-based displacement measurement. The framework starts with video capture and initial

preprocessing, after which a 2D OWT is applied to each frame to decompose the image into horizontal and vertical subbands, denoted as CH and CV. These subbands are then subjected to the Hilbert transform to generate their corresponding Hilbert pairs (CH_h and CV_h). Each original subband is subsequently combined with its Hilbert pair to construct the COWT, which provides the analytic signal required for phase extraction. Finally, the phase-based optical flow is computed from the extracted phase components (ϕ_V and ϕ_H) to derive the displacement signal for measurement.

2.1 Gaussian smoothing preprocessing

Gaussian smoothing is a widely used preprocessing method in image feature detection for reducing noise and suppressing fine details, allowing dominant structures at coarser scales to emerge [35]. It is commonly used as a preprocessing step in feature detection and displacement measurement algorithms. For example, in Canny edge detection, Gaussian smoothing is applied before gradient computation to reduce noise [36]. Additionally, in the LK optical flow method, Gaussian smoothing is used in the Gaussian pyramid to construct a multi-scale representation, enabling robust displacement measurement [11]. In the field of displacement measurement, the local area with clear edges is important for localization. Noise, characterized as a high-frequency component similar to edges, can be mistakenly identified as an edge, resulting in unreliable displacement measurement. Meanwhile, the surface with homogeneous pixel intensity imposes difficulties for displacement measurement, especially for phase-based methods [37]. With the preprocessing of Gaussian smoothing, disturbing factors such as noise and homogeneous area can be suppressed, thereby emphasizing the target's edge information [38]. In the COWT method, the Gaussian smoothing preprocessing of video is adopted to reduce the effects of noise and homogeneous patterns, while preserving key edge structures of the measurement target for further analysis. In the process of Gaussian smoothing, all frames of the video are convolved with a Gaussian operator, as shown in Eq. (1).

$$I_{gs}(x,y,t) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \otimes I(x,y,t) \tag{1}$$

where I (x, y,t) and I_{gs} (x, y,t) denote the original and Gaussian-smoothed image intensities, respectively, at spatial coordinates (x, y) in frame t; σ denotes the standard deviation of the Gaussian kernel; and ⊗denotes the two-dimensional spatial convolution operation applied independently to each video frame.

2.2 Two-dimensional orthogonal wavelet transform for video frame

In the OWT, the scale and translation parameters of the wavelet are discretized to construct orthogonal bases across scales and spatial locations, and the resulting wavelet coefficients are referred to as subbands. Owing to the orthogonality of these bases, subband aliasing is avoided and the decomposition is non-redundant [39, 40].

In the multiresolution decomposition, the mother wavelet $\psi(x)$ spans the wavelet space, while the father wavelet, i.e. the scaling function $\phi(x)$, spans the scale space. They satisfy the two-scale relations given in Eqs. (2) and (3), in which the mutually orthogonal high-pass filter $h(n)$ and low-pass filter $l(n)$ appear as the corresponding filter coefficients and n is the summation index.

$$\psi(x) = \sqrt{2}\sum_n h(n)\phi(2x-n) \tag{2}$$

$$\phi(x) = \sqrt{2}\sum_n l(n)\phi(2x-n) \tag{3}$$

Since video frames are inherently two-dimensional, the 2D OWT is employed. Its scaling function is formed as the product of two 1D scaling functions along the horizontal and vertical directions of the image, as given in Eq. (4).

$$\phi(x,y) = \phi(x)\phi(y) \tag{4}$$

The wavelet function in two dimensions can be constructed as the product of a 1D scaling function and a 1D wavelet function along either the vertical or horizontal direction. There are three wavelet functions for 2D OWT, as shown in Eqs. (5–7).

$$\psi^1(x,y) = \phi(x)\psi(y) \tag{5}$$

$$\psi^2(x,y) = \psi(x)\phi(y) \tag{6}$$

$$\psi^3(x,y) = \psi(x)\psi(y) \tag{7}$$

The decomposition process of the 2D OWT is illustrated in Fig. 2. At each scale level, the input at that level is convolved separately with the low-pass filter $l(n)$ and the high-pass filter $h(n)$ along the row and column directions, yielding four filtered results: low-pass filtering in both directions; $l(n)$ along each row with $h(n)$ along each column; $h(n)$ along each row with $l(n)$ along each column; and high-pass filtering in both directions. Each filtered result is then downsampled by two along both directions, so that it retains one quarter of the coefficients of the input at that level. The four downsampled

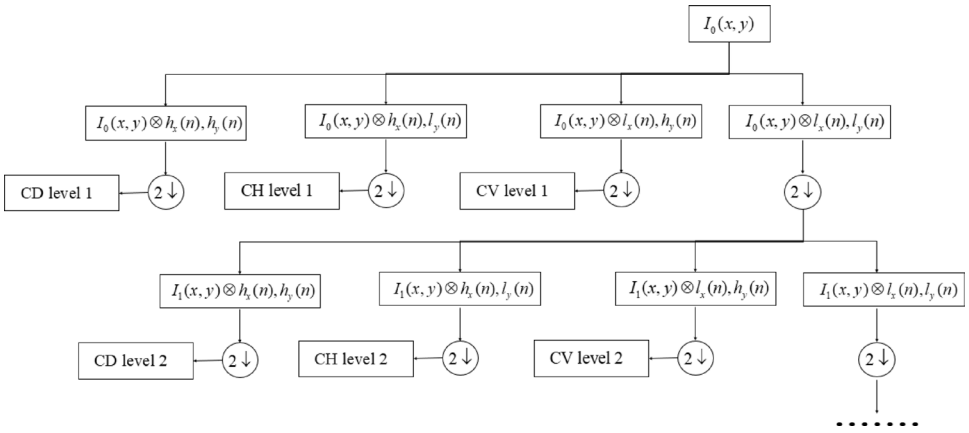


Fig. 2 The decomposition process of 2D OWT

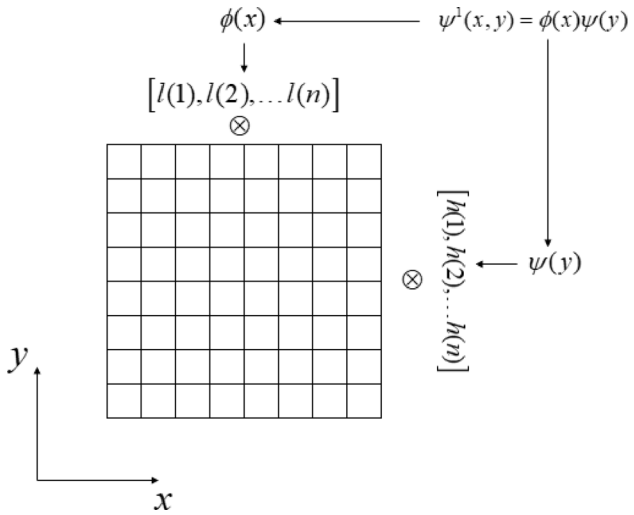


Fig. 3 The convolution process to obtain wavelet coefficients of CV

results form the subbands of this scale level: the approximation subband, which is passed to the next level, and the three detail subbands CV, CH and CD, which capture the vertical, horizontal and diagonal details of the image and correspond to the bases given in Eqs. (5–7), respectively. The convolution process for obtaining CV is illustrated in Fig. 3. Since the four subbands together contain the same number of coefficients as the input, the decomposition is non-redundant. The approximation subband is subsequently decomposed in the same manner, and the process is repeated until the maximum scale level is reached.

The Haar wavelet is adopted to perform the 2D OWT in this study [41]. As the simplest and lowest-order compactly supported orthogonal wavelet, Haar has the shortest filter support and a simple step-like structure, enabling a compact and clearly localized response to the sharp target edges commonly encountered in vision-based structural displacement measurement. In contrast, higher-order compactly supported orthogonal wavelets have more vanishing moments and generally exhibit longer, asymmetric, and more complex multi-lobed structures, which may produce multiple local extrema and more complex phase variations around the same edge. These characteristics make it difficult to establish a unique and stable correspondence between target-edge motion and the extracted phase response, thereby reducing the stability and accuracy of phase-based displacement measurement. Accordingly, Haar is selected in the present COWT framework because it enables a more direct and stable phase representation of target-edge motion.

The function of the Haar wavelet in the discrete Haar wavelet transform has the form as shown in Eq. (8):

$$\psi(x)=\begin{cases}1, & \text{for x}\in\left[0,\frac{1}{2}\right),\\ -1, & \text{for x}\in\left[\frac{1}{2},1\right),\\ 0, & \text{otherwise.}\end{cases} \tag{8}$$

The scaling function of the Haar wavelet is:

$$\phi(x)=\begin{cases}1, & \text{for x}\in[0,1),\\ 0, & \text{for x}\notin[0,1).\end{cases} \tag{9}$$

The low-pass l and high-pass filters h of Haar wavelet are given in the Eqs. (10, 11) [42].

$$\{l(0),l(1)\}=\left\{\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right\} \tag{10}$$

$$\{h(0),h(1)\}=\left\{\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right\} \tag{11}$$

2.3 The construction of COWT and phase extraction

After applying the 2D OWT to the video frame, we mainly employ the Hilbert transform to construct the COWT based on the subbands of the OWT. The transformation process is illustrated in this section.

Among the four subbands at each scale level, only CV and CH are employed to construct the COWT, since they carry the vertical and horizontal details required for bidirectional displacement measurement. For the subband CV, its Hilbert pair CV_h is obtained by applying the Hilbert transform along the vertical direction of the wavelet coefficients, as shown in Eq. (12). As defined in Eq. (13), CV_c is formed from the original OWT subband CV and its Hilbert counterpart CV_h, corresponding to the real and imaginary components, respectively. This combination yields a complex subband in the COWT and represents an analytic signal.

By performing the Hilbert transform on subband CH along the horizontal direction, its Hilbert pair CH_h is obtained, as shown in Eq. (14). Following the same procedure as for CV to construct the analytic signal, the subband CH_c is obtained, as illustrated in Eq. (15).

The subbands CV_c and CH_c together form the COWT. The name COWT reflects the way its complex subbands are constructed, which follows the same principle used to derive the CSP from the steerable pyramid, with each imaginary term supplied by the corresponding Hilbert pair [43, 44]. It can also be interpreted that the COWT is the analytic-signal version of the OWT.

$$CV_h(x_i, y, t) = \frac{1}{\pi} \int \frac{CV(x_i, \tau, t)}{y - \tau} d\tau \tag{12}$$

$$CV_c(x_i, y, t) = CV(x_i, y, t) + jCV_h(x_i, y, t) \tag{13}$$

$$CH_h(x, y_i, t) = \frac{1}{\pi} \int \frac{CH(\tau, y_i, t)}{x - \tau} d\tau \tag{14}$$

$$CH_c(x, y_i, t) = CH(x, y_i, t) + jCH_h(x, y_i, t) \tag{15}$$

After constructing the COWT, the phase information in the vertical and horizontal directions can be computed from the subbands CV_c and CH_c of the COWT using the phase angle theory, as shown in Eqs. (16) and (17). Consequently, bidirectional phase information is obtained, which enables phase-based displacement measurement in both the horizontal and vertical directions.

$$\phi_V(x_i, y, t) = \arg(CV_c(x_i, y, t)) = \arctan\left(\frac{CV_h(x_i, y, t)}{CV(x_i, y, t)}\right) \tag{16}$$

$$\phi_H(x, y_i, t) = \arg(CH_c(x, y_i, t)) = \arctan\left(\frac{CH_h(x, y_i, t)}{CH(x, y_i, t)}\right) \tag{17}$$

2.4 Displacement extraction using COWT phase information

The phase components contained in the COWT subbands CV_c and CH_c can be utilized to estimate pixel motion in the vertical and horizontal directions. Under the phase-constancy assumption, the pixel velocity and displacement between consecutive frames can then be determined [17, 45].

The procedure for deriving the component velocity in phase-based optical flow is described in [46]. The final expression for the component velocity $\vec{v_c}$ is given in Eq. (18), where ϕ_x and ϕ_y denote the spatial phase gradients in the horizontal and vertical directions, respectively, and ϕ_t represents the temporal phase gradient.

$$\vec{v_c} = -\frac{\phi_t}{({\phi_x}^2 + {\phi_y}^2)}(\phi_x, \phi_y) \tag{18}$$

The COWT method provides phase components ϕ_V and ϕ_H, which capture details in the vertical and horizontal directions, respectively. These phase components also correspond to the component velocities in the respective directions, since they share the same gradient direction as the direction in which they capture details. Therefore, Eq. (18) can be projected onto the vertical and horizontal directions for the COWT method. Finally, the component velocities in the vertical and horizontal directions can be computed from the corresponding phase information, as shown in Eqs. (19) and (20).

$$\vec{v_{cH}} = -\frac{\phi_t}{({\phi_x}^2)}(\phi_x, 0) = -\frac{\phi_t}{({\phi_x}^2)}(\frac{d\phi_H}{dx}, 0) \tag{19}$$

$$\vec{v_{cV}} = -\frac{\phi_t}{({\phi_y}^2)}(0, \phi_y) = -\frac{\phi_t}{({\phi_y}^2)}(0, \frac{d\phi_V}{dy}) \tag{20}$$

After the pixel velocity is obtained, the displacement can be determined by multiplying it by the time interval between two consecutive frames, as expressed in Eq. (21).

$$\vec{d_{cH}} = \vec{v_{cH}}\,\Delta t,\ \vec{d_{cV}} = \vec{v_{cV}}\,\Delta t \tag{21}$$

3 Verification of COWT phase-based optical flow

In this section, laboratory and outdoor tests are employed to verify the reliability and efficiency of COWT phase-based optical flow. Additionally, the accuracy and computational speed of COWT are compared with other phase extraction

methods and intensity-based methods for displacement measurement. As a result, the proposed COWT method not only reduces computational demand significantly but also provides reliable phase information for accurate displacement measurement, thereby leading to more efficient phase-based displacement measurement.

3.1 Verification test on a shaker-mounted column

An experiment using vision-based techniques was conducted to evaluate the performance of the COWT phase-based optical flow method. The experiment was designed to measure the displacement of a column fixed to the end of a shaker that generated one-dimensional reciprocating motion at adjustable speeds. The shaker stroke length was set to its minimum value of 1.18 in. (approximately 30 mm peak-to-peak). The oscillation frequency was not prescribed as a numerical input. At the start of the shaker operation, the shaker speed was manually increased stepwise until a steady operating speed was reached and maintained during the main measurement period. At the end of the shaker operation, the speed was manually reduced stepwise to zero. Therefore, the complete displacement record cannot be characterized by a single prescribed oscillation frequency. Because the motion speed could be set to a relatively low level, the video was recorded using an iPhone 11 at 24 FPS and a resolution of 720P. A Keyence LK-G3000 series laser displacement sensor was used to provide the reference displacement, with its measurement direction aligned with the motion direction of the column. The experimental layout and its schematic illustration are presented in Fig. 4a and b.

Due to the limited availability of experimental equipment, the vision-based experiment was only conducted as a one-directional motion estimation. However, as the decomposition process of the 2D OWT method is identical in both vertical and horizontal directions, this one-directional experiment sufficiently demonstrates the method's efficacy.

The optical axis of the camera was aligned perpendicular to the moving direction of the target. The camera was positioned 900 mm away from the target, with a focal length of 52 mm, corresponding to a spatial resolution of 0.402 mm/pixel. At the beginning of the video, the column was located at the frame center, as depicted in the first frame of the video shown in Fig. 5. For the COWT method, Gaussian smoothing was adopted as a preprocessing step to extract the edge feature and reduce the negative effects from the homogeneous pattern and noise [37].

The recorded video was processed to extract displacement histories using COWT and the comparative methods presented in Fig. 6. For all phase-based methods, the point of interest (POI) was placed at the center of the frame. For LK optical flow method, a point on the column edge was selected for tracking, since visual features are more distinct along the direction of motion. In the template matching, the entire column was used as the reference template.

The laser displacement sensor sampled at 1000 Hz, whereas the video was recorded at 24 Hz. For direct comparison, the two displacement histories were manually aligned in time, and the laser displacement sensor data were resampled according to the camera frame rate.

As shown in Fig. 6a, the displacement history obtained from COWT generally agrees well with the laser reference throughout the measurement period. The representative positive peak enlarged in Fig. 6b provides a clearer view of the local discrepancies among the investigated methods. While several comparative methods exhibit noticeable overestimation or underestimation of the peak displacement, the COWT result remains close to the laser reference.

Computational efficiency was evaluated on a Dell Precision 3680 workstation. All methods except RAFT were executed on a single CPU thread of an Intel Core i7-14700

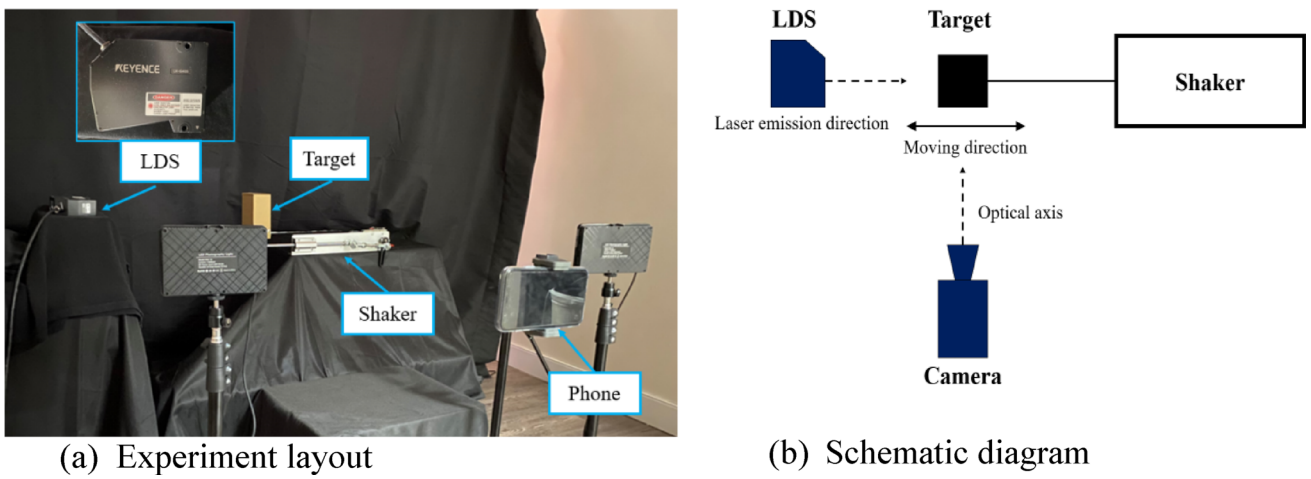


Fig. 4 Experimental setup and schematic illustration

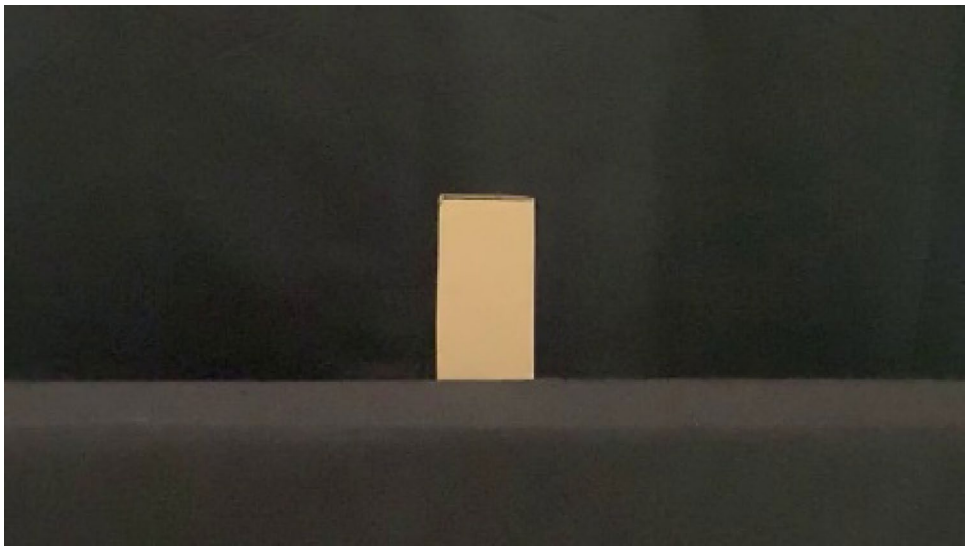


Fig. 5 The first frame of the recorded motion video

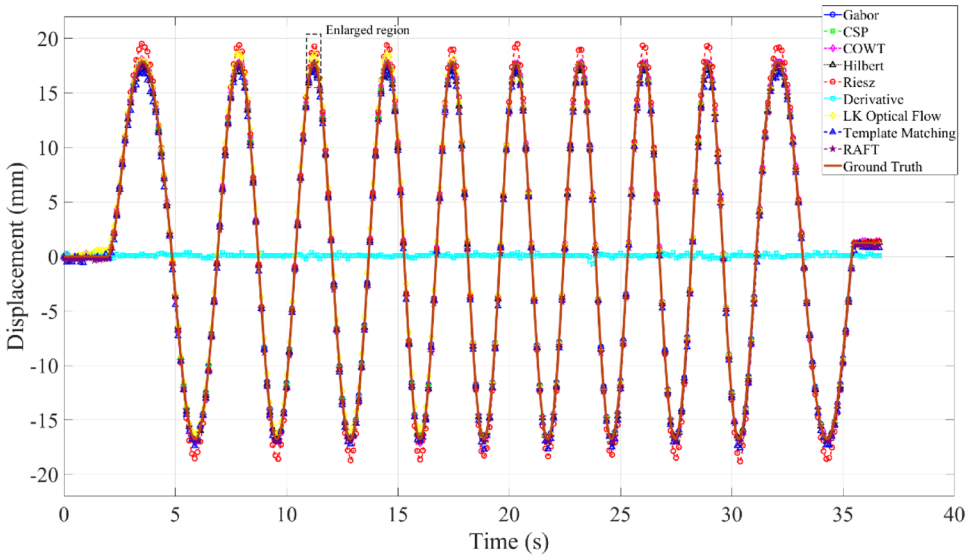
processor in a system equipped with 32 GB of DDR5 memory operating at 4400 MT/s, whereas RAFT was executed on an NVIDIA GeForce RTX 5060 Ti GPU with 16 GB of dedicated memory. Each method was independently run five times under identical software, parameter, and input-data settings. The individual computation times and corresponding mean values are reported in Table 1.

As shown in Table 1, the computation times obtained from the five repeated runs are closely distributed for all investigated methods, indicating limited run-to-run variability under identical experimental settings. Based on the mean computation times, COWT requires 13.6148 s to process the laboratory video. Among the phase-based methods executed on the CPU, COWT is approximately 13.1, 16.8, 12.3, 13.4, and 4.6 times faster than CSP, Gabor wavelet, the Hilbert-transform-enhanced method, the Riesz-transform-enhanced method, and the derivative-enhanced method, respectively. Its mean computation time is also close to those of the intensity-based methods, being approximately 6.1% lower than that of LK optical flow and essentially identical to that of template matching.

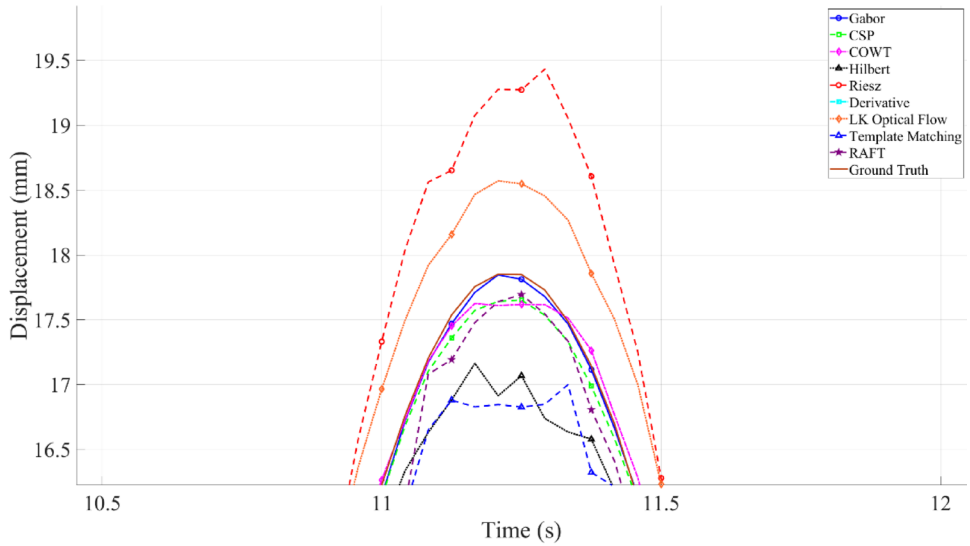
The computational advantage of COWT is mainly related to its orthogonal formulation and the non-redundant OWT decomposition framework on which it is built. By comparison, the Gabor wavelet and CSP methods require overcomplete representations, resulting in substantially longer processing times. For the Hilbert-transform-enhanced and Riesz-transform-enhanced methods, a considerable portion of the computation is associated with the forward and inverse FFT operations required for frequency-domain filtering. The derivative-enhanced method is relatively efficient because of its simple Sobel-filter convolution, but its mean computation time remains approximately 4.6 times that of COWT. The GPU-accelerated computation time of RAFT is also reported as a practical reference but is not directly compared with the CPU-based methods when evaluating relative computational efficiency.

A detailed accuracy analysis of the tested methods in the laboratory test is further conducted using both MAE and root mean square error (RMSE). MAE characterizes the average magnitude of the displacement measurement errors, whereas RMSE assigns greater weight to relatively large instantaneous errors and is therefore more sensitive to localized deviations in dynamic displacement time histories. The combined use of these two metrics provides a more comprehensive evaluation of the measurement accuracy.

As shown in Table 2, COWT achieves an MAE of 0.3147 mm and an RMSE of 0.4123 mm in the laboratory test. Both metrics produce the same ranking of the investigated methods, indicating that the accuracy comparison is consistent across the two error measures. Compared with the Hilbert-transform-, Riesz-transform-, and derivative-enhanced methods, COWT reduces the MAE/RMSE by 13.4%/8.9%, 56.5%/54.7%, and 96.8%/96.5%, respectively. Its advantage over the Hilbert- and Riesz-transform-enhanced methods is mainly related to the construction of the motion-sensitive subbands. In COWT, the decomposition level is selected to match the motion region directly, whereas the Hilbert- and Riesz-transform-enhanced approaches rely more heavily on peak-frequency selection, which can be more susceptible to noise contamination and unrelated image-frequency components. The derivative-enhanced method exhibits substantially larger MAE and RMSE values, suggesting insufficient robustness for the present measurement task. Relative to LK optical flow, template matching, and RAFT, COWT reduces the MAE/RMSE by 32.0%/23.2%, 38.5%/33.4%, and 15.8%/11.5%, respectively, demonstrating better displacement measurement accuracy than the investigated intensity-based and deep-learning-based methods in this laboratory test.



(a) Complete displacement time histories



(b) Enlarged view of the peak region

Fig. 6 Comparison of the displacement time histories obtained using the investigated methods and the laser displacement sensor in the laboratory shaker test. **a** Complete displacement time histories, **b** Enlarged view of the peak region

Table 1 Computation times (s) obtained from five repeated runs for different vision-based methods in the laboratory test

Run No.	COWT	CSP	Gabor	Hilbert	Riesz	Derivative	LK	TM	RAFT
1	13.5679	177.9354	228.4421	167.4134	180.9087	62.1752	14.6507	13.4255	233.9676
2	13.4999	179.7002	228.7139	167.8831	182.6965	61.7664	14.4533	13.5195	233.9532
3	13.7030	180.9835	231.2884	167.1361	181.8455	62.1693	14.4530	13.7728	233.9435
4	13.7863	179.0858	227.9972	166.1900	183.0834	61.9280	14.3927	13.7018	234.7272
5	13.5170	177.3990	229.0728	167.8398	182.6557	61.9668	14.2842	13.6609	234.0920
Mean	13.6148	179.0208	229.1029	167.2925	182.2380	62.0011	14.4468	13.6161	234.1367

Table 2 MAE- and RMSE-based accuracy evaluation of different methods for the laboratory test

	COWT	CSP	Gabor	Hilbert	Riesz	Derivative	LK	TM	RAFT
MAE (mm)	0.3147	0.2571	0.2269	0.3635	0.7234	9.9684	0.4627	0.5119	0.3736
RMSE (mm)	0.4123	0.3322	0.2992	0.4524	0.9094	11.6310	0.5369	0.6190	0.4658

However, COWT exhibits lower measurement accuracy than the Gabor wavelet and CSP methods. As shown in Table 2, COWT yields an MAE of 0.3147 mm and an RMSE of 0.4123 mm, compared with MAEs of 0.2269 mm and 0.2571 mm and RMSEs of 0.2992 mm and 0.3322 mm for Gabor wavelet and CSP, respectively. Both error metrics therefore produce the same accuracy ranking. The lower accuracy of COWT is mainly related to the specific filter-bank and critical-sampling design of its underlying non-redundant OWT decomposition framework. In particular, the fixed Haar filters provide limited frequency selectivity, while critical downsampling increases sensitivity to small spatial translations.

From the frequency-domain perspective, the Haar analysis filters used in COWT are fixed and divide the image frequency content into relatively broad subbands. Therefore, the frequency range used for displacement measurement can only be adjusted by selecting among discrete decomposition levels [40]. By contrast, the center frequency and bandwidth of a Gabor wavelet can be adjusted more flexibly according to the target size and motion range, allowing its response to be concentrated around the spatial frequencies most relevant to the moving target [20]. This more concentrated frequency response helps suppress background textures, image noise, and other irrelevant frequency components [46]. Consequently, the selected COWT detail subband may contain a broader mixture of target-edge, local-texture, and noise components, making its local complex response more dependent on the specific image content and resulting in a less stable phase-displacement relationship than that obtained using a well-tuned Gabor wavelet.

From the perspective of shift sensitivity, the Haar-based OWT adopted in COWT is critically sampled, with downsampling performed after filtering at each decomposition level. When a target edge undergoes a small spatial translation, its alignment with both the Haar filter support and the coefficient-sampling grid changes. As a result, the wavelet coefficients may not simply undergo an equivalent spatial translation; their amplitudes may also vary, introducing additional fluctuations into the subsequently constructed complex phase response. In comparison, Gabor wavelet is generally performed densely over the full pixel grid and therefore avoids the sampling-grid sensitivity introduced by critical downsampling [47]. CSP, by contrast, achieves shift-invariant subband representations through its overcomplete and essentially aliasing-free pyramid structure, allowing its coefficients to follow input translations more consistently [33]. Therefore, the local complex responses generated by Gabor wavelet and CSP generally vary more smoothly with target translation, allowing displacement to be represented more consistently through phase changes.

Despite these limitations, the Haar wavelet has very short spatial support and a simple step-like structure, allowing a direct and localized response to sharp edges. Therefore, the edge-related coefficients generally remain spatially concentrated and sufficiently strong after target motion. In the present laboratory test, the relatively limited frequency selectivity and shift sensitivity of the Haar decomposition mainly introduce additional phase fluctuations and a moderate loss of measurement accuracy, rather than causing measurement failure.

To evaluate the overall performance of the investigated methods from the perspectives of measurement accuracy and computational efficiency, the dimensionless Performance Efficiency Index (PEI) is defined using normalized accuracy and computational-efficiency scores. For the ith method, the normalized accuracy score A_i and computational-efficiency score E_i are defined as:

$$A_i = \frac{\mathrm{MAE}_{\min}}{\mathrm{MAE}_i}, \qquad E_i = \frac{T_{\min}}{T_i} \tag{22}$$

where MAE_i and T_i denote the MAE and computation time of the ith method, respectively, while MAE_{min} and T_{min} denote the minimum MAE and computation time among all methods compared in the same experiment. Both A_i and E_i are dimensionless and bounded between 0 and 1, with a larger value indicating better performance.

The PEI is then calculated by combining the normalized accuracy and computational-efficiency scores as:

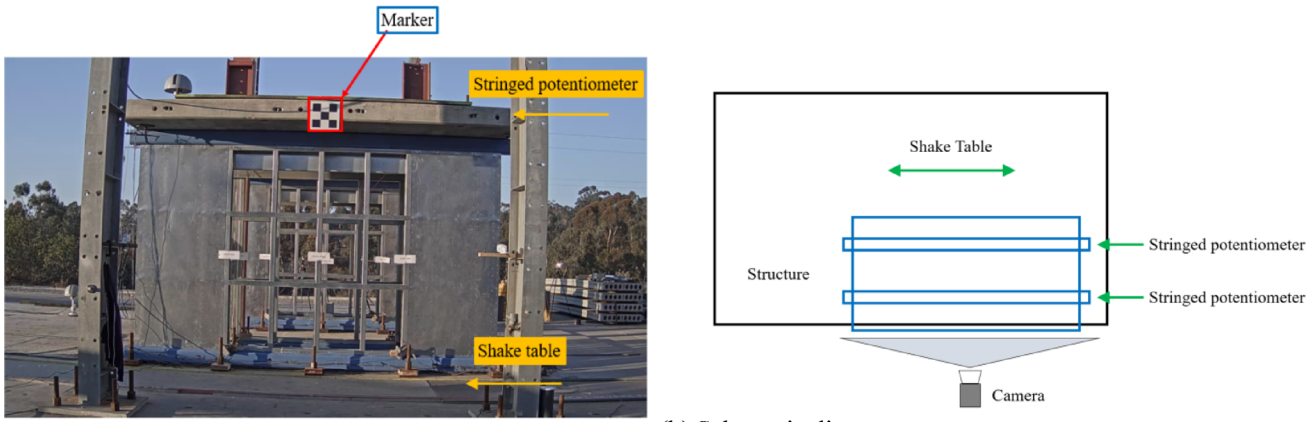
$$\mathrm{PEI}_i = A_i^{\alpha} E_i^{\beta}, \qquad \alpha + \beta = 1 \tag{23}$$

where α and β represent the relative weights assigned to measurement accuracy and computational efficiency, respectively. The balanced setting is defined as $\alpha = \beta = 0.50$, corresponding to the equally weighted geometric mean of the two normalized scores. To evaluate the sensitivity of the PEI results to the weighting choice, three settings are considered: $(\alpha,\beta) = (0.25,0.75)$, $(0.50,0.50)$, and $(0.75,0.25)$, representing efficiency-oriented, balanced, and accuracy-oriented evaluations, respectively. Since the normalization is performed separately for each experiment, the PEI values represent the relative performance of the investigated methods within the same experiment.

RAFT is excluded from the PEI evaluation because its computation time was obtained using GPU acceleration, whereas all other methods included in the PEI calculation were executed on the CPU. Directly combining GPU- and CPU-based computation times in the same PEI would introduce computational-platform effects into the efficiency score and would therefore not provide a consistent comparison.

Table 3 PEI-based overall evaluation under different weighting settings for the laboratory test

	COWT	CSP	Gabor	Hilbert	Riesz	Derivative	LK	TM
α=0.25, β=0.75	0.9215	0.1404	0.1204	0.1354	0.1069	0.1246	0.8004	0.8159
α=0.50, β=0.50	0.8491	0.2591	0.2438	0.2254	0.1531	0.0707	0.6798	0.6657
α=0.75, β=0.25	0.7824	0.4782	0.4937	0.3751	0.2191	0.0401	0.5774	0.5432



(a) Experiment layout (b) Schematic diagram

Fig. 7 Shake table test configuration for the CSW specimen. (adapted from [48]). **a** Experiment layout, **b** Schematic diagram

Nevertheless, RAFT is retained in the accuracy and computation time evaluations as a reference method.

As shown in Table 3, COWT achieves the highest PEI under all three weighting settings, with values of 0.9215, 0.8491, and 0.7824 for the efficiency-oriented, balanced, and accuracy-oriented settings, respectively. Under the efficiency-oriented setting, template matching and LK optical flow achieve the second- and third-highest PEI values because their mean computation times are close to that of COWT. Under the balanced and accuracy-oriented settings, LK optical flow ranks second, while the PEI values of Gabor wavelet and CSP increase as greater weight is assigned to measurement accuracy because of their lower MAEs. In contrast, the PEI of the derivative-enhanced method decreases substantially as the accuracy weight increases because of its relatively large measurement error. Although the rankings of several comparative methods vary with the weighting setting, COWT consistently ranks first, confirming that its favorable accuracy-efficiency balance is not dependent on a single weighting choice.

3.2 The validation through a full-scale outdoor shake table experiment

The proposed COWT method was further validated through a full-scale outdoor shake table experiment on a cold-formed steel wall (CSW) structure under earthquake excitation. This experiment was carried out at the Large High Performance Outdoor Shake Table (LHPOST) at the University of California, San Diego, USA [48].

The CSW specimens were developed from archetype designs representing four- and ten-story buildings so that the structural configuration remained consistent with the previously reported test program. Detailed information on the specimen setup, experimental program, and associated datasets is provided in the DesignSafe repository [48].

Figure 7 shows the outdoor testing setup. Earthquake input was applied to the CSW specimen through LHPOST using the CNP196 component of the 1994 Northridge Canoga Park record. A DVR camera positioned perpendicular to the excitation direction was used to capture the structural motion. The video was recorded at 1920 × 1080 resolution and 30 fps, with an image scale of 4.4418 mm/pixel based on the calibration reported by Cai et al. [49]. A marker near the top of the structure served as the tracking target, and string potentiometers mounted on the transfer beam provided reference lateral displacement data at 256 Hz.

Since real earthquakes are unpredictable and uncontrollable, conducting true field seismic experiments is practically impossible. Therefore, full-scale shake table tests serve as the most effective and realistic alternative, providing controlled yet representative seismic excitations for evaluating structural performance. Compared to small-scale or indoor experiments, full-scale outdoor shake table tests also provide a more authentic environment for vision-based methods. These tests capture natural lighting conditions, complex and dynamic backgrounds, and realistic structural responses, all of which are critical for evaluating the

robustness and applicability of vision-based measurement methods under real-world conditions.

The displacement measurement results of the COWT are presented in Fig. 8. The COWT signal closely matches the string potentiometer data, confirming the reliability of its displacement measurements.

For consistency, the outdoor test was evaluated using the comparison framework established in Sect. 3.1. The complete displacement histories are presented in Fig. 9a, while a representative positive peak is enlarged in Fig. 9b to provide a clearer comparison of the local discrepancies among the investigated methods. Among the compared approaches, the derivative-enhanced method again does not perform satisfactorily. The Hilbert-transform- and Riesz-transform-enhanced methods generally follow the trend of the reference displacement, but their estimated amplitudes show noticeable deviations. This discrepancy is likely attributable to the challenging visual conditions of the recorded video. More specifically, the tracking marker covers only a limited region in the image, whereas the surrounding CSW surface presents strong visual patterns. Under these conditions, frequency components unrelated to the true structural motion may interfere with the frequency-selection process, thereby reducing the accuracy of displacement measurement.

Table 4 presents the individual computation times and corresponding mean values obtained from five repeated runs for the outdoor test under the hardware and execution settings described in Sect. 3.1. The five computation times of each method are closely distributed, indicating limited run-to-run variability. Based on the mean computation times, COWT requires 31.5489 s to process the outdoor video. Among the phase-based methods executed on the CPU, COWT is approximately 12.7, 17.4, 9.7, 12.5, and 4.4 times faster than CSP, Gabor wavelet, the Hilbert-transform-enhanced method, the Riesz-transform-enhanced method, and the derivative-enhanced method, respectively. Its mean computation time is also close to those of the intensity-based methods, being approximately 10.9% lower than that of LK optical flow and only 1.9% higher than that of template matching. The GPU-accelerated RAFT method has a mean computation time of 539.8546 s. This value is reported as a practical reference but is not used in the direct CPU-based efficiency comparison or the PEI evaluation because RAFT and the other methods were executed on different computational platforms.

The MAE and RMSE values of the different methods are presented in Table 5. As shown in Table 5, COWT achieves an MAE of 1.8945 mm and an RMSE of 2.2085 mm in the full-scale outdoor test. The two metrics produce the same ranking of the investigated methods, indicating that the accuracy comparison is consistent across different error measures. Compared with the Riesz-transform-enhanced method, which achieves the lowest MAE and RMSE among the other recently developed phase extraction methods, COWT reduces the MAE and RMSE by 60.6% and 63.0%, respectively. Relative to the Hilbert-transform-enhanced and derivative-enhanced methods, the corresponding reductions are 70.5%/82.3% and 87.8%/93.2%, respectively. COWT also outperforms LK optical flow, template matching, and RAFT. In particular, the considerably larger increases from MAE to RMSE observed for the Hilbert-transform-enhanced, derivative-enhanced, and RAFT methods indicate the presence of relatively large instantaneous deviations in their displacement histories. By contrast, the smaller RMSE-to-MAE increase of COWT suggests a more consistent error distribution throughout the measured response.

However, the Gabor wavelet and CSP methods retain lower errors, with MAE/RMSE values of 1.1404/1.5396 mm

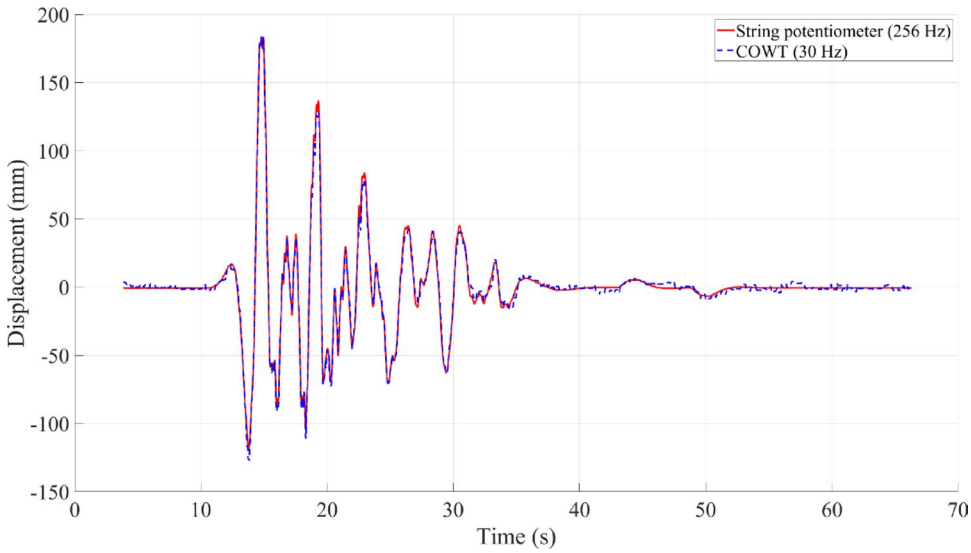
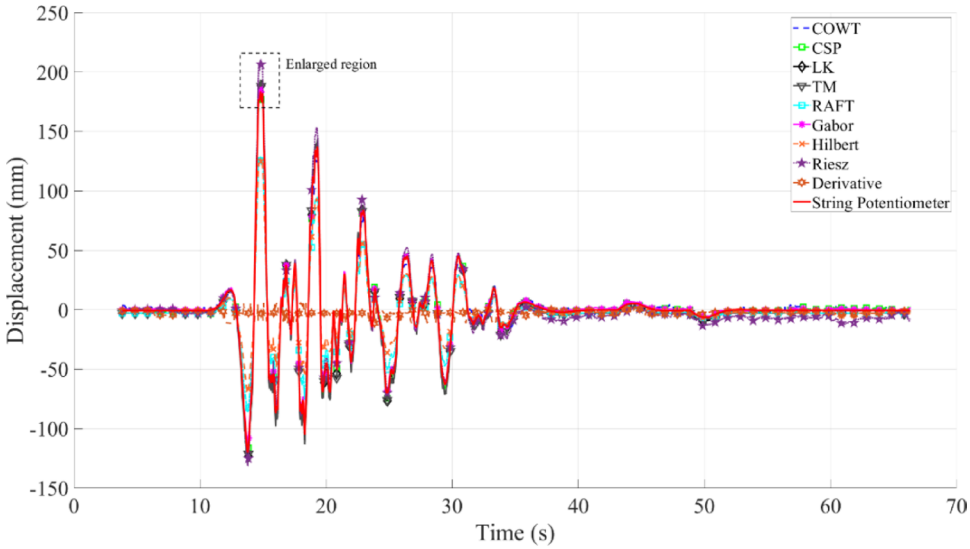
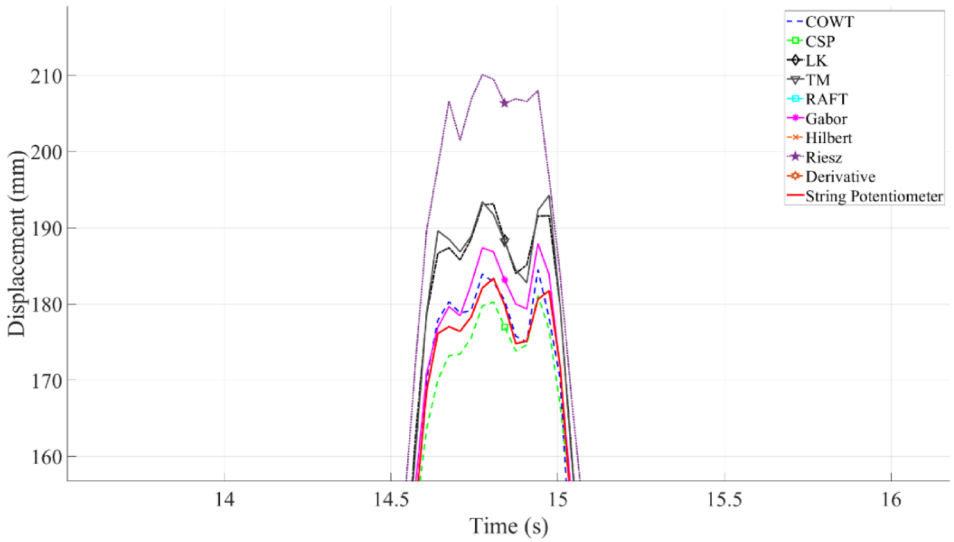


Fig. 8 Comparison of the lateral displacement time histories of the CSW specimen obtained using COWT and the string potentiometer under the CNP196 earthquake excitation



(a) Complete displacement time histories



(b) Enlarged view of the peak region

Fig. 9 Comparison of the lateral displacement time histories of the CSW specimen obtained using the investigated methods and the string potentiometer. **a** Complete displacement time histories, **b** Enlarged view of the peak region

Table 4 Computation times (s) obtained from five repeated runs for different methods in the CSW shake table experiment

Run No.	COWT	CSP	Gabor	Hilbert	Riesz	Derivative	LK	TM	RAFT
1	31.3186	401.8812	548.9593	305.4621	393.8886	139.7189	35.5046	31.5827	540.8521
2	31.7647	400.6495	546.1027	303.0511	395.6036	139.9084	35.3787	30.6717	537.9318
3	31.6762	399.9000	549.3839	305.3793	396.4711	139.8737	35.3217	31.3258	539.8376
4	31.4910	401.0457	547.3664	303.1052	394.5463	138.3468	35.1671	30.9246	542.5104
5	31.4938	400.3594	550.1364	305.2556	395.7244	138.8940	35.5762	30.3280	538.1411
Mean	31.5489	400.7672	548.3897	304.4507	395.2468	139.3484	35.3897	30.9666	539.8546

Table 5 MAE- and RMSE-based error assessment of different methods for the CSW shake table study

	COWT	CSP	Gabor	Hilbert	Riesz	Derivative	LK	TM	RAFT
MAE (mm)	1.8945	1.3397	1.1404	6.4154	4.8023	15.5743	2.6180	2.8131	5.8719
RMSE (mm)	2.2085	1.6944	1.5396	12.4529	5.9643	32.4581	3.0593	3.3981	10.2075

Table 6 PEI-based overall evaluation under different weighting settings for the CSW shake table experiment

	COWT	CSP	Gabor	Hilbert	Riesz	Derivative	LK	TM
$\alpha=0.25, \beta=0.75$	0.8686	0.1408	0.1158	0.1169	0.1034	0.1684	0.7350	0.7979
$\alpha=0.50, \beta=0.50$	0.7687	0.2565	0.2376	0.1345	0.1364	0.1276	0.6174	0.6367
$\alpha=0.75, \beta=0.25$	0.6802	0.4672	0.4875	0.1546	0.1800	0.0966	0.5186	0.5080

and 1.3397/1.6944 mm, respectively, compared with 1.8945/2.2085 mm for COWT. This result is consistent with the filter-level analysis presented in Sect. 3.1. In the outdoor video, the tracking marker occupies only a limited region of the image, whereas the surrounding CSW surface contains strong structural textures and other unrelated spatial-frequency components. Under these conditions, the relatively broad dyadic frequency band of the Haar decomposition may contain more irrelevant image components, making the local complex response more dependent on the surrounding image content. In addition, the shift sensitivity associated with critical downsampling may introduce additional phase fluctuations when the target undergoes small spatial translations. By comparison, the stronger frequency selectivity of the Gabor wavelet and the shift-invariant representation of CSP provide more stable phase responses under target translation. These factors help explain why Gabor wavelet and CSP maintain higher measurement accuracy than COWT in the outdoor test.

The same normalized PEI formulation and the three weighting settings defined in Sect. 3.1 are applied to the outdoor test to evaluate the sensitivity of the accuracy-efficiency comparison to different weighting choices.

As shown in Table 6, COWT again achieves the highest PEI under all three weighting settings, with values of 0.8686, 0.7687, and 0.6802 for the efficiency-oriented, balanced, and accuracy-oriented settings, respectively. Template matching and LK optical flow occupy the next two positions under the efficiency-oriented and balanced settings because of their relatively short mean computation times. Under the accuracy-oriented setting, LK optical flow ranks second and template matching ranks third. As greater weight is assigned to measurement accuracy, the PEI values and relative rankings of Gabor wavelet and CSP improve because they produce the lowest MAEs in the outdoor test. However, their substantially longer computation times prevent them from surpassing COWT. The derivative-enhanced method performs relatively better under the efficiency-oriented setting because of its moderate computational cost, but its PEI decreases considerably as the accuracy weight increases because of its large measurement error. Although the rankings among some comparative methods vary with the weighting setting, COWT consistently maintains the highest PEI, confirming its favorable accuracy-efficiency balance under different evaluation preferences in the full-scale outdoor test.

Although the laboratory and outdoor tests demonstrate the effectiveness and computational efficiency of COWT, the present experimental validation remains limited in scope. The laboratory test considered a single target undergoing one-dimensional reciprocating motion over the investigated speed range. The outdoor test was based on one earthquake excitation, with displacement evaluated at a single tracking marker. Consequently, the reported results should be interpreted as validation under the tested conditions rather than as a comprehensive assessment across all possible monitoring scenarios. Further validation involving different target characteristics, multidirectional motions, broader speed ranges, multiple earthquake excitations, and multiple tracking locations would help assess the general applicability of the method.

4 Conclusion

This study proposes a complex orthogonal wavelet transform (COWT) phase extraction method built on a non-redundant OWT decomposition framework to reduce the computational cost of phase-based displacement measurement. A laboratory shaker test and a full-scale outdoor shake table test were employed to evaluate its measurement accuracy and computational efficiency.

In the laboratory test, COWT achieved an MAE of 0.3147 mm and an RMSE of 0.4123 mm, with a mean computation time of 13.6148 s. In the outdoor test, the corresponding values were 1.8945 mm, 2.2085 mm, and 31.5489 s, respectively. Specifically, COWT was 16.8 and 13.1 times faster than the Gabor wavelet and CSP methods, respectively, in the laboratory test, and 17.4 and 12.7 times faster, respectively, in the outdoor test. It was also 4.4–13.4 times faster than the other recently developed phase extraction methods. These substantial computational advantages support its potential for real-time displacement measurement. Relative to the most accurate of the other recently developed phase extraction methods, COWT reduced the MAE and RMSE by 13.4% and 8.9%, respectively, in the laboratory test, and by 60.6% and 63.0%, respectively, in the outdoor test. Moreover, COWT achieved the highest PEI under all three weighting settings in both experiments, demonstrating a favorable balance between measurement accuracy and computational efficiency.

Although COWT exhibited lower measurement accuracy than the Gabor wavelet and CSP methods, its substantially lower computational cost and reliable performance make it a promising alternative for efficient phase-based displacement measurement. Future work will focus on improving the frequency selectivity and shift invariance of the COWT framework to further enhance measurement accuracy.

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Dong C-Z, Catbas FN (2021) A review of computer vision–based structural health monitoring at local and global levels. Struct Health Monit 20:692–743. https://doi.org/10.1177/1475921720935585
2. Saxena SC, Seksena SBL (1989) A self-compensated smart LVDT transducer. IEEE Trans Instrum Meas 38:748–753. https://doi.org/10.1109/19.32186
3. Gomez F, Park J-W, Spencer BF (2018) Reference-free structural dynamic displacement estimation method, Struct. Control Health Monit 25:e2209. https://doi.org/10.1002/stc.2209
4. Feng D, Feng MQ (2016) Vision-based multipoint displacement measurement for structural health monitoring: vision-based displacement measurement for SHM. Struct Control Health Monit 23:876–890. https://doi.org/10.1002/stc.1819
5. Chen Y, Joffre D, Avitabile P (2018) Underwater dynamic response at limited points expanded to full-field strain response. J Vib Acoust 140:051016. https://doi.org/10.1115/1.4039800
6. Xu Y, Brownjohn JMW (2018) Review of machine-vision based methodologies for displacement measurement in civil structures. J Civ Struct Health Monit 8:91–110. https://doi.org/10.1007/s13349-017-0261-4
7. Feng D, Feng MQ (2018) Computer vision for SHM of civil infrastructure: from dynamic response measurement to damage detection—a review. Eng Struct 156:105–117. https://doi.org/10.1016/j.engstruct.2017.11.018
8. Ghorbani R, Matta F, Sutton MA (2015) Full-field deformation measurement and crack mapping on confined masonry walls using digital image correlation. Exp Mech 55:227–243. https://doi.org/10.1007/s11340-014-9906-y
9. Baqersad J, Poozesh P, Niezrecki C, Avitabile P (2017) Photogrammetry and optical methods in structural dynamics—a review. Mech Syst Signal Process 86:17–34. https://doi.org/10.1016/j.ymssp.2016.02.011
10. Horn BKP, Schunck BG (1981) Determining optical flow. Artif Intell 17:185–203. https://doi.org/10.1016/0004-3702(81)90024-2
11. Lucas BD, Kanade T (1981) An iterative image registration technique with an application to stereo vision. In: 7th international joint conference on Artificial intelligence, Vancouver, pp 674–679. https://hal.science/hal-03697340
12. Javh J, Slavič J, Boltežar M (2017) The subpixel resolution of optical-flow-based modal analysis. Mech Syst Signal Process 88:89–99. https://doi.org/10.1016/j.ymssp.2016.11.009
13. Bhowmick S, Nagarajaiah S, Lai Z (2020) Measurement of full-field displacement time history of a vibrating continuous edge from video. Mech Syst Signal Process 144:106847. https://doi.org/10.1016/j.ymssp.2020.106847
14. Lydon D, Lydon M, Taylor S, Del Rincon JM, Hester D, Brownjohn J (2019) Development and field testing of a vision-based displacement system using a low cost wireless action camera. Mech Syst Signal Process 121:343–358. https://doi.org/10.1016/j.ymssp.2018.11.015
15. Lydon D, Lydon M, Del Rincon JM, Taylor SE, Robinson D, O'Brien E, Catbas FN (2018) Development and field testing of a time-synchronized system for multi-point displacement calculation using low-cost wireless vision-based sensors. IEEE Sens J 18:9744–9754. https://doi.org/10.1109/JSEN.2018.2853646
16. Teed Z, Deng J (2020) Recurrent all-pairs field transforms for optical flow. In: Vedaldi A, Bischof H, Brox T, Frahm J-M (eds) Computer Vision – ECCV 2020. Springer, Cham, pp 402–419. https://doi.org/10.1007/978-3-030-58536-5_24.
17. Fleet DJ, Jepson AD (1990) Computation of component image velocity from local phase information. Int J Comput Vis 5:77–104. https://doi.org/10.1007/BF00056772
18. Barron JL, Fleet DJ, Beauchemin SS (1994) Performance of optical flow techniques. Int J Comput Vis 12:43–77
19. Diamond DH, Heyns PS, Oberholster AJ (2017) Accuracy evaluation of sub-pixel structural vibration measurements through optical flow analysis of a video sequence. Measurement 95:166–172. https://doi.org/10.1016/j.measurement.2016.10.021
20. Miao Y, Jeon JY, Kong Y, Park G (2022) Phase-based displacement measurement on a straight edge using an optimal complex Gabor filter. Mech Syst Signal Process 164:108224. https://doi.org/10.1016/j.ymssp.2021.108224
21. Wadhwa N, Rubinstein M, Durand F, Freeman WT (2013) Phase-based video motion processing. ACM Trans Graph 32:1–10. https://doi.org/10.1145/2461912.2461966
22. Chen JG, Wadhwa N, Cha Y-J, Durand F, Freeman WT, Buyukozturk O (2015) Modal identification of simple structures with high-speed video using motion magnification. J Sound Vib 345:58–71. https://doi.org/10.1016/j.jsv.2015.01.024
23. Chen JG, Wadhwa N, Cha Y-J, Durand F, Freeman WT, Buyukozturk O (2014) Structural modal identification through high speed camera video: motion magnification. In: De Clerck J (ed) Topics in modal analysis, vol 7. Springer, Cham, pp 191–197. https://doi.org/10.1007/978-3-319-04753-9_19.
24. Sarrafi A, Mao Z, Niezrecki C, Poozesh P (2018) Vibration-based damage detection in wind turbine blades using phase-based motion estimation and motion magnification. J Sound Vib 421:300–318. https://doi.org/10.1016/j.jsv.2018.01.050
25. Collier S, Dare T (2022) Accuracy of phase-based optical flow for vibration extraction. J Sound Vib 535:117112. https://doi.org/10.1016/j.jsv.2022.117112
26. Yang Y, Dorn C, Mancini T, Talken Z, Kenyon G, Farrar C, Mascareñas D (2017) Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification. Mech Syst Signal Process 85:567–590. https://doi.org/10.1016/j.ymssp.2016.08.041
27. Liu G, Li MZ, Mao Z, Yang QS (2022) Structural motion estimation via Hilbert transform enhanced phase-based video processing. Mech Syst Signal Process 166:108418. https://doi.org/10.1016/j.ymssp.2021.108418
28. Li MZ, Liu G, Mao Z, Yang QS, Gu JW (2023) Two-dimensional motion estimation using phase-based image processing with Riesz transform. Mech Syst Signal Process 188:110044. https://doi.org/10.1016/j.ymssp.2022.110044

29. Liu S, Yu L, Niu W, Wang J, Zhong Z, Huang J, Shan M (2024) Fast and accurate visual vibration measurement via derivative-enhanced phase-based optical flow. Mech Syst Signal Process 209:111089. https://doi.org/10.1016/j.ymssp.2023.111089
30. Luan L, Zheng J, Wang ML, Yang Y, Rizzo P, Sun H (2021) Extracting full-field subpixel structural displacements from videos via deep learning. J Sound Vib 505:116142. https://doi.org/10.1016/j.jsv.2021.116142
31. Yang J, Tse P (2021) Sparse representation of complex steerable pyramid for machine fault diagnosis by using non-contact video motion to replace conventional accelerometers. Measurement 175:109104. https://doi.org/10.1016/j.measurement.2021.109104
32. Simoncelli EP, Freeman WT (1995) The steerable pyramid: a flexible architecture for multi-scale derivative computation. In: Proceedings of the IEEE International Conference on Image Processing (ICIP), Washington, DC, pp 444–447. https://doi.org/10.1109/ICIP.1995.537667
33. Simoncelli EP, Freeman WT, Adelson EH, Heeger DJ (1992) Shiftable multiscale transforms. IEEE Trans Inf Theory 38:587–607. https://doi.org/10.1109/18.119725
34. Polikar R (1996) The Wavelet Tutorial: The Engineer's Ultimate Guide to Wavelet Analysis. https://users.rowan.edu/~polikar/WTtutorial.html. Accessed 03 Aug 2026
35. Basu M (2002) Gaussian-based edge-detection methods-a survey. IEEE Trans Syst Man Cybern Part C Appl Rev 32:252–260. https://doi.org/10.1109/TSMCC.2002.804448
36. Canny J (1986) A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell PAMI 8:679–698. https://doi.org/10.1109/TPAMI.1986.4767851
37. Miao Y, Kong Y, Nam H, Lee S, Park G (2023) Phase-based vibration imaging for structural dynamics applications: Marker-free full-field displacement measurements with confidence measures. Mech Syst Signal Process 198:110418. https://doi.org/10.1016/j.ymssp.2023.110418
38. Torre V, Poggio TA, Detection OE (1986) IEEE Trans Pattern Anal Mach Intell PAMI 8:147–163. https://doi.org/10.1109/TPAMI.1986.4767769
39. Simpson DM (1993) An introduction to the discrete orthogonal wavelet transform. Rev Bras Eng 9:57–81
40. Mallat SG (1989) A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell 11:674–693. https://doi.org/10.1109/34.192463
41. Haar A (1909) Zur theorie der orthogonalen funktionensysteme. Georg-August-Universitat, Gottingen
42. Porwik P, Lisowska A (2004) The Haar–Wavelet Transform in digital image processing. Its status and achievements, p 13
43. Portilla J, Simoncelli EP (2000) A parametric texture model based on joint statistics of complex wavelet coefficients. Int J Comput Vis 40:49–70
44. Wadhwa N, Rubinstein M, Durand F, Freeman WT (2014) Riesz pyramids for fast phase-based video magnification. In: 2014 IEEE International Conference on Computational Photography (ICCP), Santa Clara, pp 1–10 https://doi.org/10.1109/ICCPHOT.2014.6831820
45. Gautama T, Van Hulle MA (2002) A phase-based approach to the estimation of the optical flow field using spatial filtering. IEEE Trans Neural Netw 13:1127–1136. https://doi.org/10.1109/TNN.2002.1031944
46. Ai S, Dong C, Mei Q (2025) The optimal scale level of complex steerable pyramid for phase-based motion estimation under different motion ranges and target sizes. Measurement 253:117440. https://doi.org/10.1016/j.measurement.2025.117440
47. Movellan JR (2002) Tutorial on Gabor filters. Open Source Doc 40:1–23
48. Singh A, Hutchinson T, Wang X, Zhang Z, Schafer B, Derveni F, Castaneda H, Peterman K (2023) Wall line tests: phase 1—shake table tests. In: CFS-NHERI: Seismic resiliency of repetitively framed mid-rise cold-formed steel buildings. https://doi.org/10.17603/DS2-MVJ8-8386
49. Cai E, Zhang Y, Lu X, Ji X, Gao X, Hou J, Shi J, Guo W (2024) Estimating structural motions in extreme environmental conditions—a dynamic correlation filter based computer vision approach. Mech Syst Signal Process 215:111398. https://doi.org/10.1016/j.ymssp.2024.111398

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.