Mechanical Systems and Signal Processing 85 (2017) 567–590

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification

Yongchao Yang [a,*], Charles Dorn [b], Tyler Mancini [c], Zachary Talken [d], Garrett Kenyon [e], Charles Farrar [a], David Mascareñas [a]

[a] Los Alamos National Lab, Engineering Institute, PO Box 1663, MS T001, Los Alamos, NM 87545, USA
[b] Department of Engineering Physics, University of Wisconsin, Madison, WI 53706, USA
[c] Department of Aerospace Engineering and Engineering Mechanics, University of Texas, Austin, TX 78712, USA
[d] Department of Mechanical and Aerospace Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA
[e] Los Alamos National Lab, Applied Modern Physics, PO Box 1663, MS D454, Los Alamos, NM 87545, USA

ARTICLE INFO

Article history:
Received 22 January 2016
Received in revised form
5 August 2016
Accepted 28 August 2016

Keywords:
Operational modal analysis
Non-contact measurements
Video processing
Blind source separation
Motion magnification

ABSTRACT

Experimental or operational modal analysis traditionally requires physically-attached wired or wireless sensors for vibration measurement of structures. This instrumentation can result in mass-loading on lightweight structures, and is costly and time-consuming to install and maintain on large civil structures, especially for long-term applications (e.g., structural health monitoring) that require significant maintenance for cabling (wired sensors) or periodic replacement of the energy supply (wireless sensors). Moreover, these sensors are typically placed at a limited number of discrete locations, providing low spatial sensing resolution that is hardly sufficient for modal-based damage localization, or model correlation and updating for larger-scale structures. Non-contact measurement methods such as scanning laser vibrometers provide high-resolution sensing capacity without the mass-loading effect; however, they make sequential measurements that require considerable acquisition time. As an alternative non-contact method, digital video cameras are relatively low-cost, agile, and provide high spatial resolution, simultaneous, measurements. Combined with vision based algorithms (e.g., image correlation, optical flow), video camera based measurements have been successfully used for vibration measurements and subsequent modal analysis, based on techniques such as the digital image correlation (DIC) and the point-tracking. However, they typically require speckle pattern or high-contrast markers to be placed on the surface of structures, which poses challenges when the measurement area is large or inaccessible. This work explores advanced computer vision and video processing algorithms to develop a novel video measurement and vision-based operational (output-only) modal analysis method that alleviate the need of structural surface preparation associated with existing vision-based methods and can be implemented in a relatively efficient and autonomous manner with little user supervision and calibration. First a multi-scale image processing method is applied on the frames of the video of a vibrating structure to extract the local pixel phases that encode local structural vibration, establishing a full-field spatiotemporal motion matrix. Then a high-spatial dimensional, yet low-modal-dimensional, over-complete model is used to represent the extracted full-field motion matrix using modal superposition, which is physically connected and manipulated by a family of unsupervised learning models and techniques, respectively. Thus, the proposed method is able to

* Corresponding author.
E-mail addresses: yyang@lanl.gov, yangyongchaohit@gmail.com (Y. Yang), charlesjdorn@gmail.com (C. Dorn), tylermancini@gmail.com (T. Mancini), zrtalken@gmail.com (Z. Talken), gkenyon@lanl.gov (G. Kenyon), farrar@lanl.gov (C. Farrar), dmascarenas@lanl.gov (D. Mascareñas).

http://dx.doi.org/10.1016/j.ymssp.2016.08.041
0888-3270/Published by Elsevier Ltd.

blindly extract modal frequencies, damping ratios, and full-field (as many points as the pixel number of the video frame) mode shapes from line of sight video measurements of the structure. The method is validated by laboratory experiments on a bench-scale building structure and a cantilever beam. Its ability for output (video measurements)-only identification and visualization of the weakly-excited mode is demonstrated and several issues with its implementation are discussed.

Published by Elsevier Ltd.

1. Introduction

Structural dynamics based vibration methods, centering on modal analysis, are among the most widely used techniques for health monitoring and model characterization and updating of civil, mechanical, and aerospace structures [1]. Experimental and operational modal analysis are two general classes of methods used to identify the dynamic properties of structures from measured data. Experimental modal analysis refers to methods that identify such properties when a measured input to the structure is available while operational modal analysis relies only on measured response data. Traditional experimental and operational modal analysis require physically-attached wired or wireless sensors, such as accelerometers that are affixed to the structure, for vibration measurements. While these sensors are reliable, they could result in mass-loading on lightweight structures, and their installation on larger structures is a costly, time and labor-consuming process. For long-term applications such as structural health monitoring (SHM), significant maintenance is required for cabling (wired sensors) or to assure adequate battery power (wireless sensors). Furthermore, these sensors only provide sparse, discrete point-wise measurements, yielding low spatial sensing resolution that is usually inadequate for modal-based damage detection, model correlation and updating of larger-scale structures. For example, a comparative study [2] showed that the spatial resolution of the sensor measurement critically limits the effectiveness of a family of popular mode shape or mode shape curvature based damage detection and localization methods. Additionally, abundant high-resolution experimental or operational measurement data and modal information (e.g., resonant frequencies and mode shapes) are extremely valuable for more accurately correlating and updating highly refined and detailed finite element models [3,4].

Non-contact vibration measurement techniques, such as displacement measurements made with scanning laser vibrometers [5–7], provide high spatial resolution sensing capacity without the need of sensor installed on the structures or inducing the mass-loading effect. However, these measurement devices are relatively expensive and perform measurements sequentially, which could be time and labor tedious when the desired sensing areas are large. As an alternative non-contact method, digital video cameras are relatively low-cost, agile, and provide simultaneous measurements with very high spatial resolution. Combined with image processing algorithms (e.g., image correlation [8] , optical flow [9]), video camera based measurements have been successfully used for vibration measurement of various types of structures [10–18]. Recently, digital image correlation (DIC) and the 3-dimensional point-tracking techniques have been introduced in experimental modal analysis [19–24]. A noticeable advantage of these methods is that high-resolution mode shapes can be obtained in a relatively efficient manner. However, these methods typically require speckle pattern or high-contrast markers to be placed on the surface of structures for deformation computation based on image intensity correlation or feature point tracking, which raises the surface preparation and target installation issue especially when the measurement area is large or inaccessible.

In order for modal analysis methods based on video camera measurements to gain wide acceptance, it is most desirable to develop methods that utilize the video measurements only without additional structural surface preparation. Lately, a video processing based method using phase-based optical flow computation [25,26] and video motion magnification technique [27] has been proposed for operational modal analysis [28] , with the advantage of providing high resolution mode shapes without the use of paints or markers on the structure's surface. In addition, it enables convenient visualization of the vibration mode. However, this procedure of the proposal depends on several user input parameters and supervision that are not suited for efficient and automated implementation in operational modal analysis. Furthermore, this video motion magnification based technique is unable to handle closely-spaced modes and the interpretation of the motion magnification technique in modal parameters identification is unclear.

By modifying the phase-based video motion magnification framework, this study aims to develop a novel output-only (video measurements) modal analysis algorithm that requires no structural surface preparation and that can be implemented in a relatively efficient and autonomous manner. Using a multi-scale pyramid decomposition and representation method [29] and the unsupervised learning approach, blind source separation (BBS) [30], to extract, model, and manipulate the full-field spatiotemporal pixel phases that encode the local structural vibration in the video measurement only, the proposed method is able to blindly extract resonant frequencies, damping ratios, and high-resolution mode shapes from line of sight video measurements of the structure. The proposed method requires little user inputs or supervision to blindly identify modes using BSS, which is able to handle closely-spaced modes. In addition, the motivation and interpretation of motion magnification is presented for identification and visualization of the weakly-excited mode with subtle vibration

Nomenclature

C	Damping matrix
I	Image intensity
$\hat{I}$	"Magnified" image intensity
$\tilde{\hat{I}}$	Noisy "magnified" image intensity
M	Mass matrix
N	Number of pixels
R	Filter response
$\tilde{R}$	Noisy filter response
$\hat{R}$	"Magnified" filter response
$\tilde{\hat{R}}$	Noisy "magnified" filter response
T	Number of temporal samples
U	Left-singular vector matrix
V	Right-singular vector matrix
W	De-mixing matrix
j	Imaginary operator
k	Stiffness element
m	Mass element
n	Number of modes
q	Modal coordinate
$\hat{q}$	"Magnified" modal coordinate
r	Rank of matrix
t	Time index
u	Left-singular vector
v	Right-singular vector
x	Pixel coordinate
y	Displacements
$\bar{y}$	Reconstructed displacements
z	Noise magnitude
Υ	Mixing matrix of principal components
Σ	Singular value matrix
Φ	Mode shape matrix
α	Magnification factor
β	Attenuation factor
γ	Mixing vector of principal component
δ	Motion
δ'	Centered phase
$\hat{\delta}'$	"Magnified" centered phase
η	Principal components
$\hat{\eta}$	"Magnified" principal components
μ	Mass-proportional damping coefficient
ρ	Amplitude
σ	Single value
φ	Mode shape vector
ψ	Phase
ω	Spatial scale
ω_0	Spatial frequency
$\mathbb{N}$	Filter response of noise

motion, which is quite common and challenging in operational modal identification, especially in the presence of noise. Laboratory experiments are to validate the developed method. The results are compared with those using accelerometers and several issues in implementation are discussed.

The remainder of the paper is organized as follows. Section 2 introduces the proposed operational modal analysis method using BSS in the phase-based video motion magnification framework and discusses its distinctions from existing work. Section 3 validates this method when applied to video measurements from the laboratory experiments on a three-story building structure and a cantilever beam structure, showing its implementation procedures and discussing the results. Section 4 draws the conclusions and points out some future work.

2. Theory

Structural vibration can be measured into video records containing frames with temporally displaced or translated image intensity $I(x+\delta(x,t))$ where x is the pixel coordinate and $\delta(x,t)$ is the spatially local, temporally varying motion. Among the widely-used image processing techniques for motion estimation, basically, digital image correlation and other similar template matching methods [8] estimate $\delta(x,t)$ by maximizing the image intensity correlation between the current frame $I(x+\delta(x,t))$ and the reference frame $I(x)$, while the optical flow techniques [9] estimate the motion field by regularizing the current and reference frames under the optical flow or images brightness consistency constraint. Both methods are based on image intensity. Just like oscillating motion can be characterized by its magnitude and phase using Fourier transform, structural vibration motion is encoded in the local amplitudes and local phases of the image measurements, which can be extracted by multi-scale, localized filters. Especially, phase information has been shown to provide a better approximation of the motion than amplitudes (related to the intensity strength) and be relatively insensitive to changes in illumination, perspective, and surface conditions [25] , making it preferable over the original image intensity information in many real-world applications. In this section, a novel phase-based output-only (video measurement) modal identification method is derived to extract modal frequencies, damping ratios, and high-resolution mode shapes of structures. The robustness of the technique to noise is also presented.

2.1. Local phase-based vibration motion representation

Because $\delta(x,t)$ encoded in $I(x+\delta(x,t))$ is spatially local, transforms (filters) that are spatially multi-scale and localized are needed to extract the local phases. A popular such transform is wavelet transform; however, its subband representation is aliased, and therefore not suited for motion representation and estimation. The complex steerable pyramid [29] has multi-

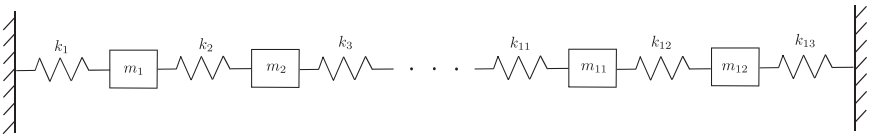

Fig. 1. A spring-mass damped numerical model with 12 DOF's.

scale, localized, non-aliasing subbands and is used to extract the local amplitudes and local phases. Applying the complex steerable pyramid filters [27,31] (complex Gabor filters that are sinusoids (spatial frequency ω_0) windowed by a Gaussian envelop leading to finite spatial support) to *each* frame (N pixels) of the video (T time samples) to build a pyramid and then collapsing it yields

$$I\big(x+\delta(x,t)\big)=\sum_{\omega=-\infty}^{\infty}R_\omega(x,t)=\sum_{\omega=-\infty}^{\infty}\rho_\omega(x,t)e^{j2\pi\omega_0(x+\delta(x,t))} \tag{1}$$

where $R_\omega(x,t)$ is the subband representation (filter response) on the spatial scale ω

$$R_\omega(x,t)=\rho_\omega(x,t)e^{j2\pi\omega_0(x+\delta(x,t))} \tag{2}$$

with the local amplitude $\rho_\omega(x,t)$ (corresponding to edge strength) and the local phase $\psi(x,t)=2\pi\omega_0\big(x+\delta(x,t)\big)=2\pi\omega_0 x+2\pi\omega_0\delta(x,t)$ that encodes the temporal vibration motion of the structure.

2.2. Modal superposition of the phase-based vibration motion

After applying the complex steerable pyramids to the frames of the video, on each spatial scale ω, the obtained local phase $\psi(x,t)$ contains the vibration motion $\delta(x,t)$, upon which output-only modal identification can be conducted. Let $\delta'(x,t)=2\pi\omega_0\delta(x,t)$ after removing the (temporal) mean, $2\pi\omega_0 x$, of the phase $\psi(x,t)$. The ($2\pi\omega_0$-factored) vibration motion $\delta'(x,t)$ can be expressed by the modal superposition as a linear combination of the modal responses

$$\delta'(x,t)=\Phi(x)q(t)=\sum_{i=1}^{n}\varphi_i(x)q_i(t) \tag{3}$$

where n is the mode number, $\Phi\in\mathbb{R}^{N\times n}$ (on Scale 1 the pixel dimension is N) is the mode shape matrix with $\varphi_i(x)$ as the ith mode shape, and $q\in\mathbb{R}^{n\times T}$ is the modal response vector with $q_i(t)$ as the ith modal coordinate. For output-only modal identification, both of Φ and $q(t)$ are to be identified from the obtained local phase $\delta'(x,t)$ only. Because the pixel (spatial) dimension of $\delta'(x,t)$, N, is much higher than the modal dimension (number), n, i.e., $N\gg n$, Eq. (3) is a high-spatial-dimensional, low-modal-dimensional *over-complete* model representing the spatiotemporal $\delta'(x,t)$ and the output-only modal identification problem above can not be directly solved. A family of unsupervised machine learning approach, known as blind source separation (BSS), is adopted to explicitly model and tackle this over-complete output-only modal identification problem because of their efficient and effective identification capability that requires little user supervision. The proposed BSS-based method is a two-step formulation: dimension reduction and modal separation, as detailed in the following.

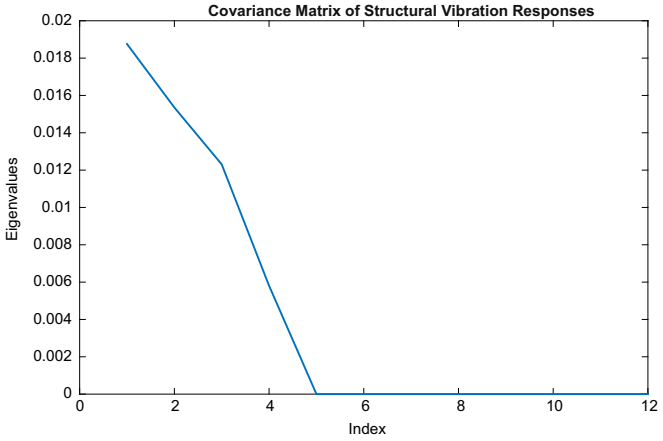

Fig. 2. The eigenvalue distribution of the covariance matrix of the structural responses of the numerical model with over-complete setting.

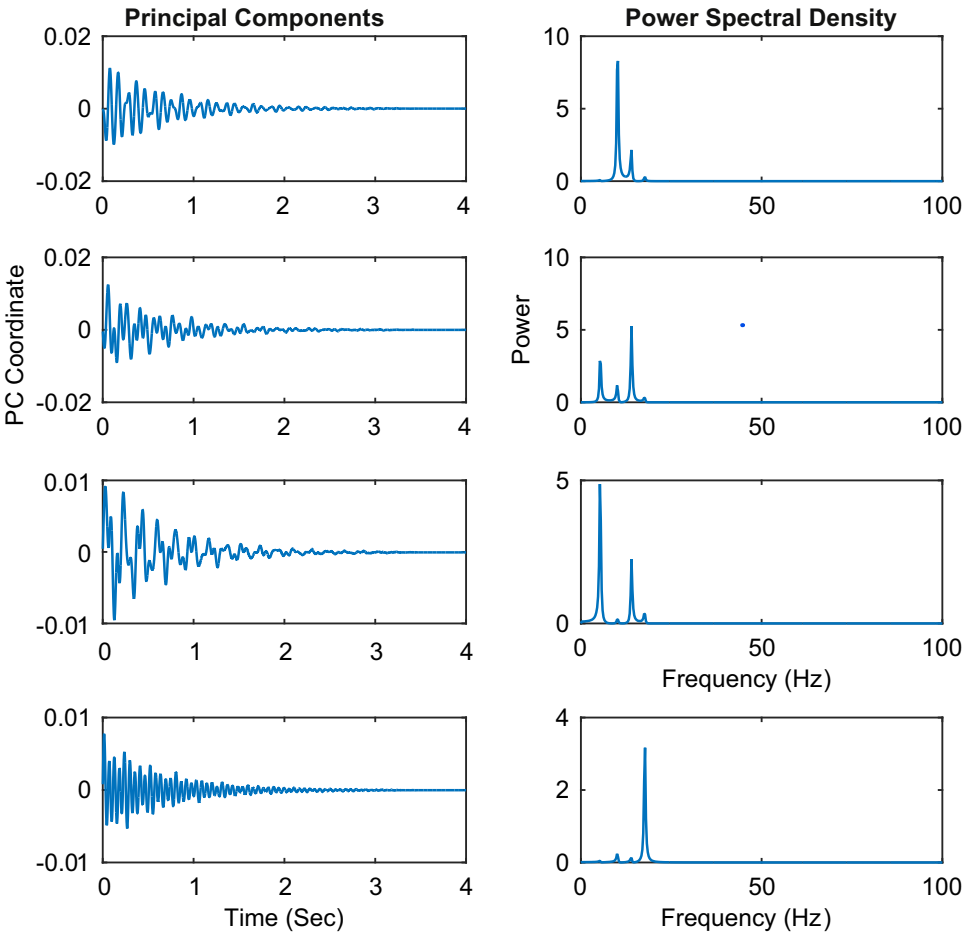

Fig. 3. The principal components and their power spectral density (PSD) of the structural responses of the numerical model.

2.2.1. Dimension reduction by principal component analysis

Dimension reduction of the over-complete system described by Eq. (3) is accomplished by PCA. As detailed in the following, the principal components are closely related to the modal components and PCA has the advantage of preserving the high spatial resolution vibration information.

Starting with the singular value decomposition (SVD) of the motion matrix $\delta' \in \mathbb{R}^{N\times T}$ (assuming $N > T$ without loss of generality)

$$\delta' = U\Sigma V^* = \sum_{i=1}^{n} \sigma_i u_i v_i^* \tag{4}$$

where $\Sigma \in \mathbb{R}^{N\times T}$ is a diagonal matrix containing T diagonal elements σ_i as the ith singular value ($\sigma_1 \geq \cdots \geq \sigma_i \geq \cdots \geq \sigma_T \geq 0$), and $U = [u_1, \ldots, u_N] \in \mathbb{R}^{N\times N}$ and $V = [v_1, \ldots, v_T] \in \mathbb{R}^{T\times T}$ are the matrices of the left- and right-singular vectors obtained by the eigenvalue decomposition (EVD) of the covariance matrices of δ'

$$\delta'\delta'^* = U\Sigma^2 U^* \tag{5}$$

$$\delta'^*\delta' = V\Sigma^2 V^* \tag{6}$$

The rank of δ' is r if the number of non-zero singular values is r ($\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_T = 0$).

It is known that the ith singular value σ_i is related to the energy projected onto the ith principal direction (vector) u_i of δ'. For structural dynamic analysis, it has been shown that for an undamped or very lightly damped structure, if its mass matrix is proportional to the identity matrix (i.e., uniform mass distribution), then the principal directions will converge to the mode shape direction [32] with the corresponding singular values indicating their participating energy in the structural vibration responses δ'. In other words, the structure's active modes, under broadband excitation, are projected onto the r principal components. An empirical, but usually true, observation is that typically there are only few dominant active modes present in the structural vibration responses, so the rank of δ', r, which is approximately the number of active singular values ($\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} \approx \cdots \approx \sigma_T \approx 0$), is very small compared to the spatial dimension of δ, N; i.e., $r \ll N$. Therefore PCA is able to

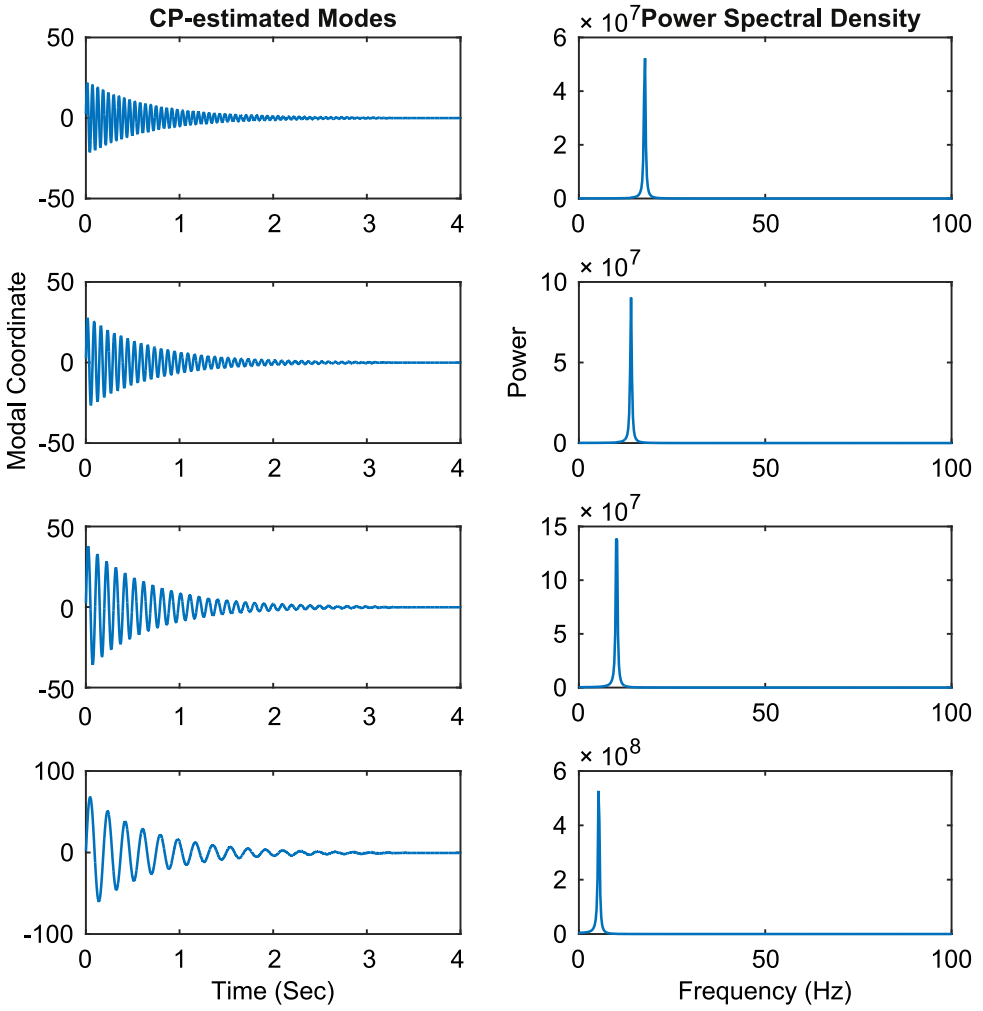

Fig. 4. The estimated modal coordinates by the complexity pursuit (CP) algorithm applied on the active principal components of the structural responses of the numerical model.

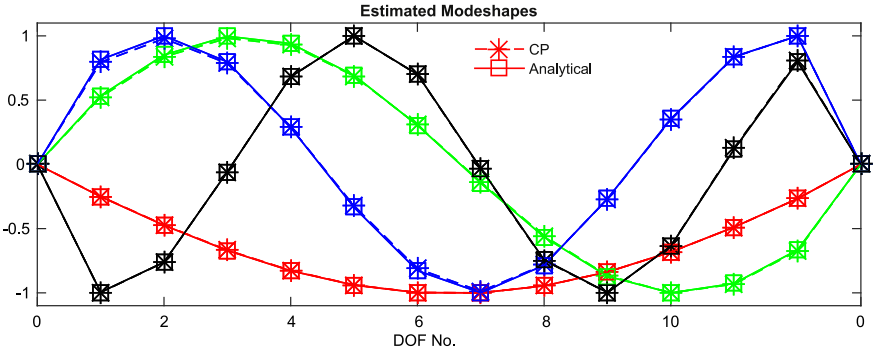

Fig. 5. The mode shapes estimated by the PCA-CP algorithm compared with the analytical ones of the numerical model. (Red: Mode 1; Green: Mode 2; Blue: Mode 3; Black: Mode 4; Star markers: Estimated; Square markers: Analytical.). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

conduct significant dimension reduction on δ' by linearly projecting most energy (modal components) of δ' onto a small number of principal components

$$\eta = U_r^* \delta' \tag{7}$$

where $U_r = \left[u_1, \ldots, u_r \right] \in \mathbb{R}^{N \times r}$ is the first r ($\ll N$) columns of U, and the ith row of the resultant $\eta = [\eta_1, \ldots, \eta_r]^* \in \mathbb{R}^{r \times T}$, η_i, is the

Table 1
Modal parameters of an over-complete numerical system estimated by PCA-CP.

Mode	Frequency (Hz)		Damping ratio (%)		MAC
	Theoretical	Estimated	Theoretical	Estimated	
1	5.3503	5.3448	4.46	4.48	1.0000
2	10.1317	10.1269	2.36	2.36	0.9999
3	13.9983	13.9920	1.71	1.71	0.9999
4	17.6846	17.6738	1.35	1.35	1.0000

ith principal component of δ'. Note that the above linear, invertible, PCA transform preserves the high spatial resolution of the low-rank $\delta' \in \mathbb{R}^{N\times T}$ by its inverse transform (projection) of $\eta \in \mathbb{R}^{r\times T}$ back to

$$\delta' \approx U_r \eta \tag{8}$$

2.2.2. Blind identification of vibration modes

As per Eqs. (4)–(7), the active mode components within $\delta' \in \mathbb{R}^{N\times T}$ are projected to $\eta \in \mathbb{R}^{r\times T}$ (so $r \approx n$) by PCA. Since the uniform mass distribution assumption [32] is generally not satisfied, each principal component of $\eta \in \mathbb{R}^{r\times T}$ is still a mixture of modes and can also be expressed as a linear combination of the modal coordinates

$$\eta(t) = \Upsilon q(t) = \sum_{i=1}^{r} \gamma_i q_i(t) \tag{9}$$

Substituting Eq. (3) and Eq. (9) into Eq. (8) yields

$$\sum_{i=1}^{n} \varphi_i q_i(t) = \delta' \approx U_r \eta = U_r \Upsilon q(t) = U_r \sum_{i=1}^{r} \gamma_i q_i(t) = \sum_{i=1}^{r} (U_r \gamma_i) q_i(t) \tag{10}$$

Comparing the two ends of Eq. (10) with $r \approx n$, it is approximately true that

$$\varphi_i = U_r \gamma_i \tag{11}$$

Recent research in structural dynamics [33–42] has established that there is a one-to-one mapping between the modal superposition model and the linear mixture model of the BSS techniques, which can perform very efficient output-only modal identification. While a family of BSS techniques is suited for output-only modal identification, previous studies [39,44] show that the complexity pursuit (CP) algorithm [43] is efficient and able to identify closely-spaced and highly-damped modes with little expert supervision and parameter adjustment and is therefore adopted in this study. Therefore using CP, $\eta(t)$ can be directly de-coupled or separated into individual modal coordinates

$$q(t) = \mathrm{W}\eta(t) \tag{12}$$

simultaneously yielding the de-mixing or de-coupled matrix $\mathrm{W} \in \mathbb{R}^{r\times r}$ and the modal coordinates $q(t)$. Comparing Eq. (9) and Eq. (12), $\Upsilon \in \mathbb{R}^{r\times r}$ can then be estimated as the inverse of W by

$$\Upsilon = \mathrm{W}^{-1} \tag{13}$$

Therefore mode shapes $\varphi_i(x)$, with very high spatial resolution (every pixel measurement on the structure), can be estimated following Eq. (11) and Eq. (13). Modal frequencies and damping ratios can be estimated from the obtained modal coordinates $q_i(t)$ by Fourier transform and logarithm decrement techniques, respectively, or using Hilbert transform (HT) based single-degree-of-freedom (SDOF) analysis methods [45] .

2.2.3. Numerical validation of PCA-CP for output-only modal identification of over-complete system

Before introducing the video motion magnification technique for identifying and analyzing vibration modes in the next section, numerical simulations are presented to validate the proposed PCA-CP method for output-only modal identification in an over-complete setting. A 12-DOF mass-spring-damper system (Fig. 1) was constructed with the following parameters (same with those in Ref. [39]): diagonal mass matrix M with m_1=2, m_2=…m_{11}=1, m_{12}=3, stiffness k_1=…=k_{13}=20000, and proportional damping matrix $C = \mu M$ with $\mu = 3$. Free vibration was simulated by an initial unit velocity at the 12th DOF. 10-second 12-channel structural displacement responses $y(t)$ are computed by the Newmark-beta algorithm with a sampling frequency of 1000 Hz.

The purpose of this example is to intentionally simulate the over-complete scenario in modeling and processing the very high spatial dimension pixel motion (of thousands or even millions of pixels) but very low modal dimension (very few modes are excited out in the operational vibration in practice). Here it is simulated that only 4 modes are contained in 12-DOF responses to emulate an over-complete scenario, which was implemented as follows. First, 12 theoretical modal

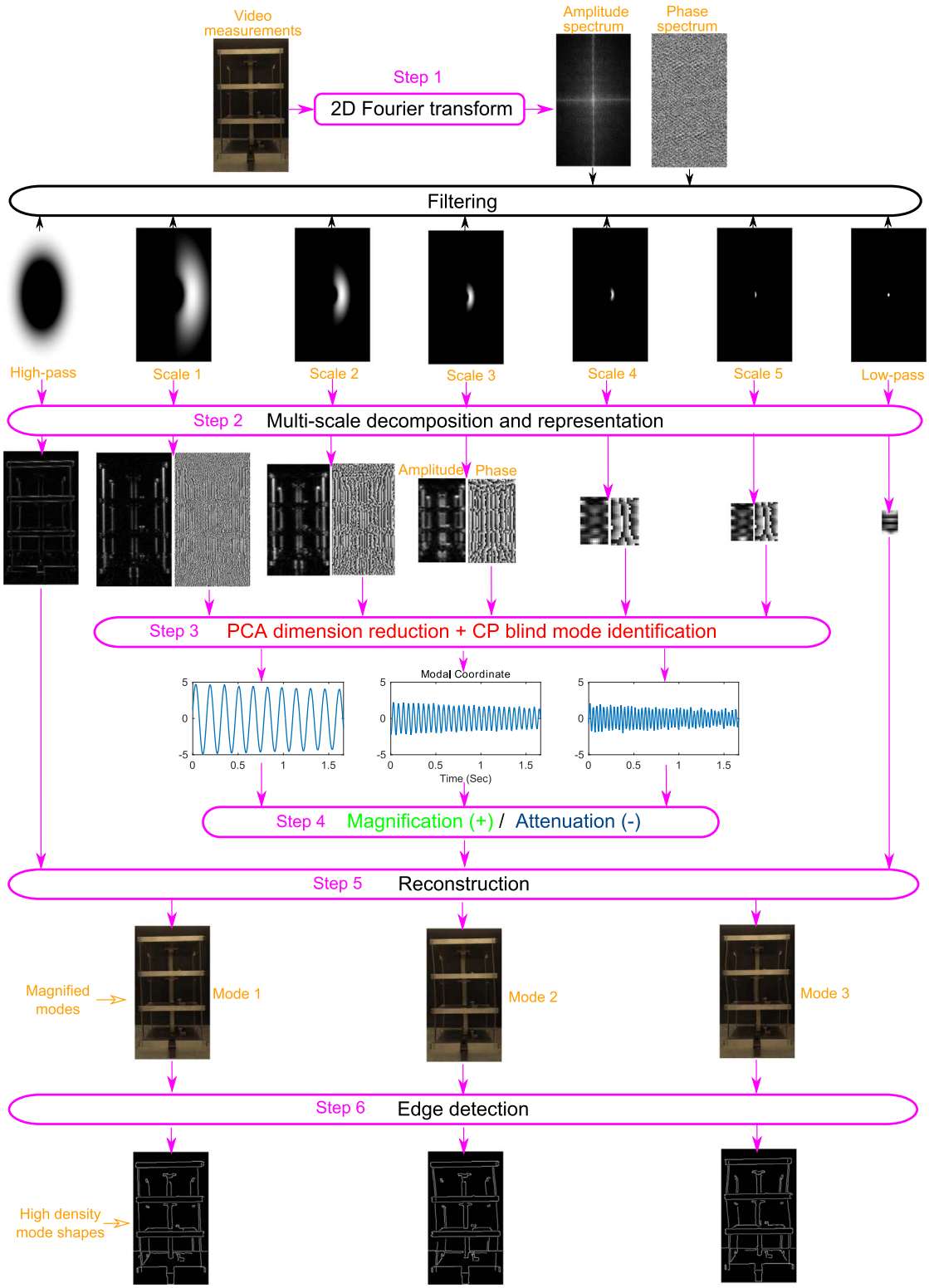

Fig. 6. The flowchart of the proposed video measurement and processing based operational modal analysis method.

coordinates $q(t)$ were obtained by multiplying $y(t)$ by the inverse of the theoretical mode shape matrix Φ; that is, $q(t) = \Phi^{-1}y(t)$. Then, only the first 4 modes were used to reconstruct a new set of structural responses $\bar{y}(t) = \sum_{i=1}^{4} \varphi_i q_i(t)$. Therefore, the new $\bar{y}(t)$ contained 12-channel structural responses but only 4 modes, which is over-complete. The PCA-CP

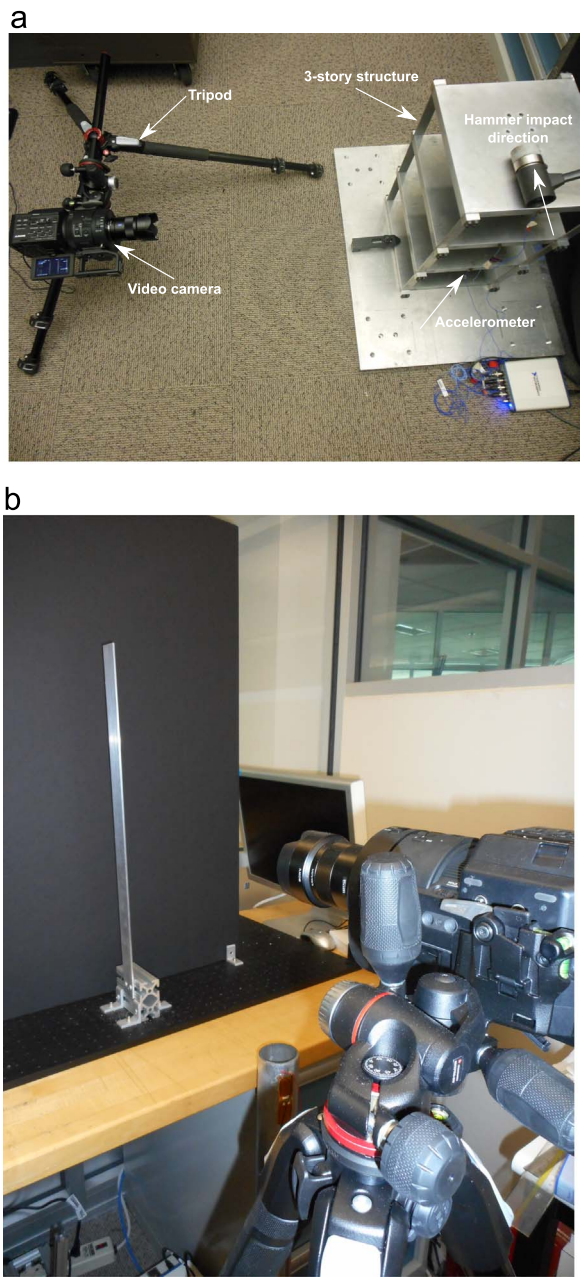

Fig. 7. The experimental setup to perform video measurement of **(a)** a three-story building structure and **(b)** a beam structure, both excited by horizontal impact.

algorithm was then applied on $\bar{y}(t)$ as per Eqs. (4)–(13). Fig. 2 shows that there are only 4 active non-zero eigenvalues (squares of the singular values), indicating most of the modal components contained in the 12-channel $\bar{y}(t)$ have been projected on to the first 4 principal components, which is depicted in Fig. 3. It has also been shown in Fig. 3 that while each of the 4 principal components contains one dominant mode of the four modes, respectively, it is still a mixture of the 4 modes. After applying CP on these 4 principal components according to Eq. (12), they are completely separated into four individual modal coordinates (shown in Fig. 4), upon which HT-based SDOF analysis methods are performed to extract the modal frequencies and damping ratios. The mode shapes were estimated according to Eq. (11) and shown in Fig. 5. Table 1 lists the estimated modal parameters which are compared with the theoretical ones, indicating they match each other very well and validating the effectiveness of the PCA-CP algorithm for output-only modal identification of over-complete systems.

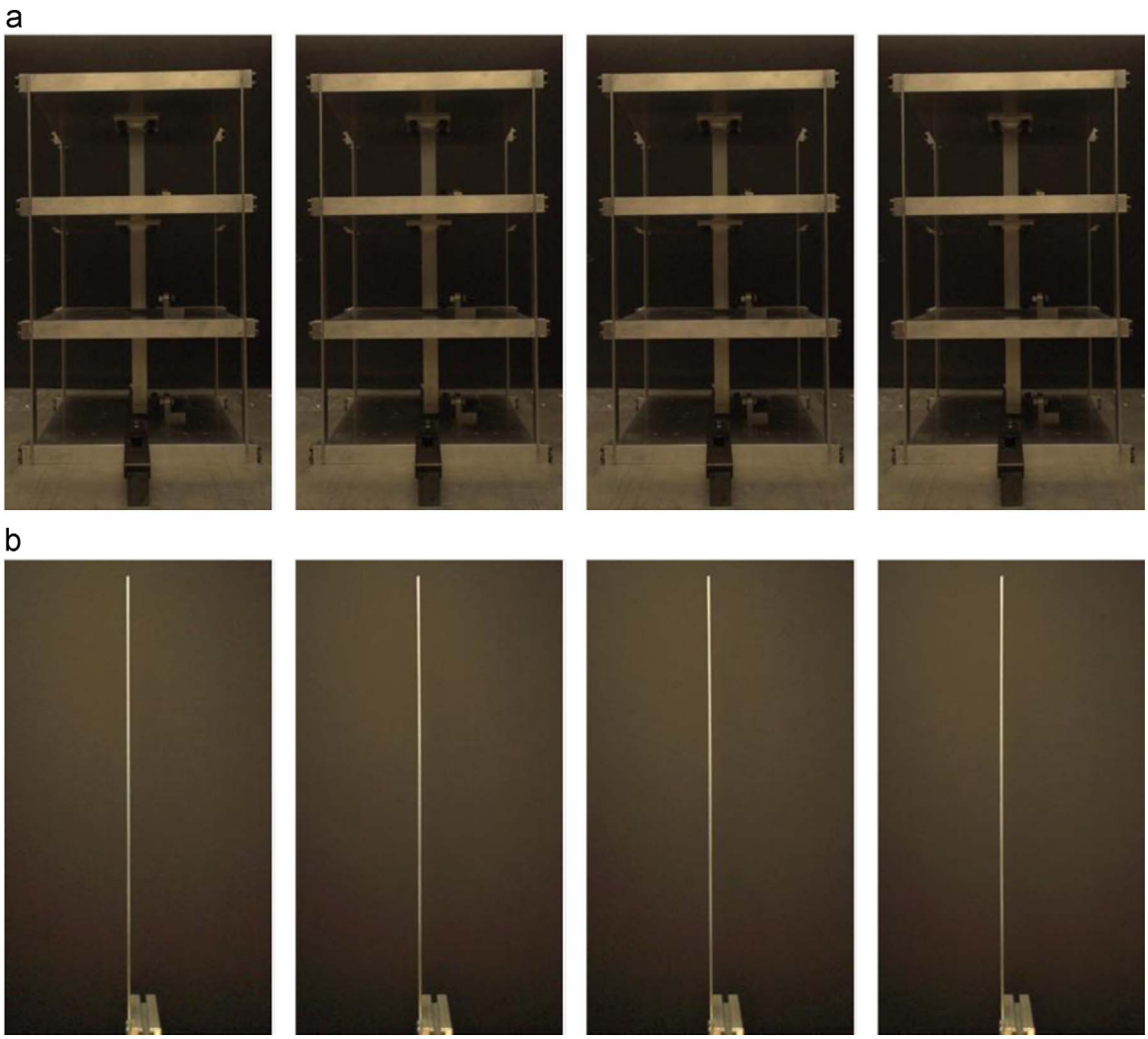

Fig. 8. (a) F**irst row:** A few temporal frames from the video measurements of the three-story **building structure** (linked to **Video 1**); **(b)** Second row: those of the **beam structure** (linked to **Video 10**).

2.3. Motion magnification of vibration modes

From Section 2.2, modal parameters (frequencies, high-resolution mode shapes, and damping ratios) can be blindly extracted. Since the proposed method does not make an assumption of the number of modes, it is able to identify as many modes as present in the vibration response measurements. However, in operational or output-only modal identification, a noticeable challenging is that some modes, usually higher modes, are very difficult to excite out; as a result, these modes are only weakly or subtly present in the structural vibration responses δ'. While modal frequencies (and damping ratios) of these weakly-excited modes (with very subtle vibration motion) may still be estimated reasonably accurately, the estimation accuracy of their mode shapes—very high resolution (involving a large amount of pixels) in the proposed method—is significantly undermined and easily contaminated by noise. This issue can be addressed by further manipulating the modal contents in the video measurements using the phase-based motion magnification technique [27] , which is robust to noise; it is detailed in the following.

2.3.1. Magnification of individual vibration mode

Following Section 2.2.1, after extracting the individual modal coordinates $q_i(t)$, motion magnification or attenuation can then be applied on them. Magnifying the ith mode is implemented as follows. First multiplying $q_i(t)$ by a magnification factor α and the rest $q_l(t)$ $(l \in [1, r] \neq i)$ by an attenuation factor $\beta = -1$ and substituting into Eq. (9) yields

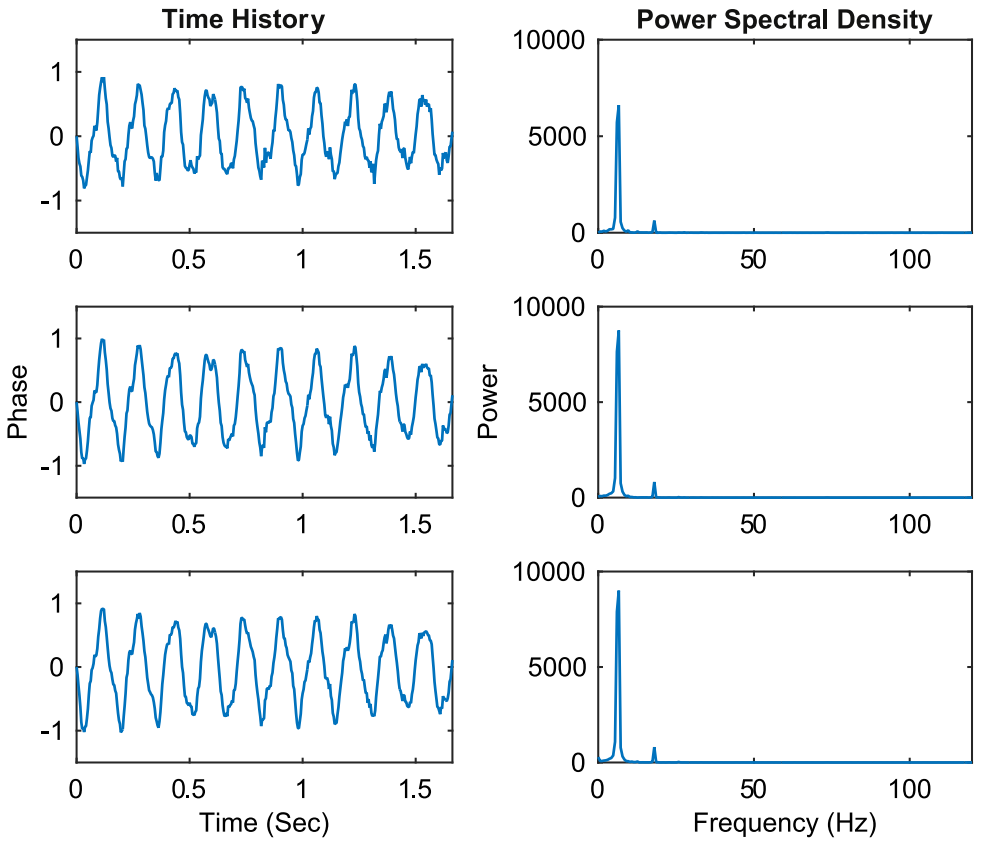

Fig. 9. The time histories and their power spectral density (PSD) of the extracted local phases of some three pixels on spatial Scale 1 from the video measurements of the **building structure**.

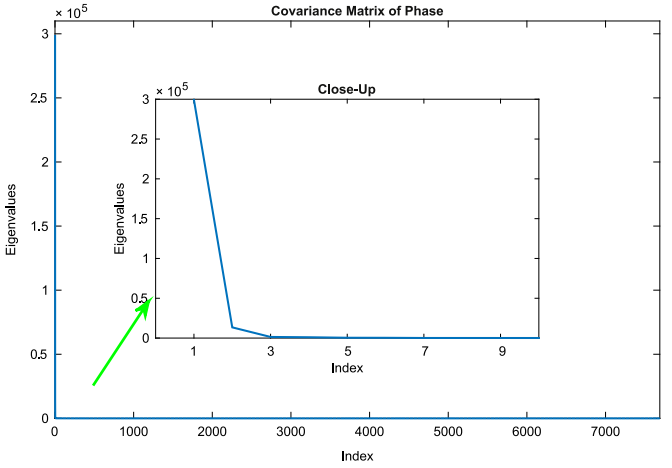

Fig. 10. The eigenvalue distribution of the covariance matrix of the spatio-temporal phases of the video (on Scale 1) of the **building structure**.

$$\hat{\eta}(t)=\Upsilon\hat{q}(t)=\gamma_i(\alpha q_i(t))+\sum_{l=1,l\neq i}^{r}\gamma_l(\beta q_l(t))=\alpha\gamma_i q_i(t)+\sum_{l=1,l\neq i}^{r}-\gamma_l q_l(t) \tag{14}$$

where "^" denotes "magnified". Substituting Eq. (9) into Eq. (14) yields

$$\hat{\eta}(t)=\alpha\gamma_i q_i(t)+\sum_{l=1,l\neq i}^{r}-\gamma_l q_l(t)=\alpha\gamma_i q_i(t)+\gamma_i q_i(t)-\eta=(1+\alpha)\gamma_i q_i(t)-\eta \tag{15}$$

Further reconstructing $\hat{\delta}'$ using Eq. (8) and Eq. (15)

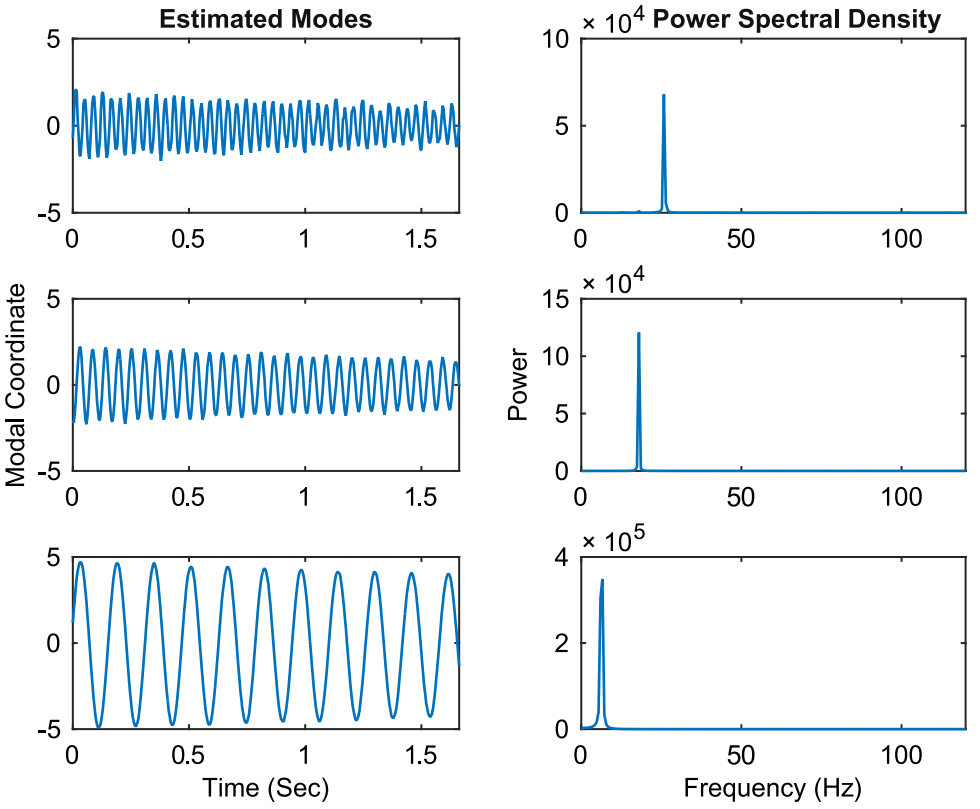

Fig. 11. The estimated modal coordinates by the complexity pursuit (CP) algorithm applied on the active principal components of the spatio-temporal pixel phases of the video of the **building structure**.

Table 2
The modal parameters of the 3-story structure estimated from the video measurements compared to those estimated from the accelerometers.

	Frequency (Hz)		Damping ratio (%)	
Mode	Accelerometer	Video	Accelerometer	Video
1	6.27	6.31	0.28	0.27
2	17.88	17.92	0.30	0.30
3	25.82	25.92	0.25	0.25

$$\hat{\delta}' \approx U_r\hat{\eta} = U_r\left[(1+\alpha)\gamma_i q_i(t) - \eta\right] = (1+\alpha)(U_r\gamma_i)q_i(t) - U_r\eta \tag{16}$$

Substituting Eq. (11) and Eq. (8) into Eq. (16) yields

$$\hat{\delta}' \approx (1+\alpha)(U_r\gamma_i)q_i(t) - U_r\eta = (1+\alpha)\varphi_i q_i(t) - \delta' \tag{17}$$

The effect of motion magnification is finally seen via phase manipulation (shift) by multiplying the original subband representation $R_\omega(x,t)$ by $e^{j\hat{\delta}'}$ and following Eq. (2) and Eq. (17)

$$\hat{R}_\omega(x,t) = R_\omega(x,t)e^{j\hat{\delta}'} = \rho_\omega(x,t)e^{j(2\pi\omega_0 x+\delta'+\hat{\delta}')} = \rho_\omega(x,t)e^{j\{2\pi\omega_0[x+(1+\alpha)\varphi_i q_i(t)/(2\pi\omega_0)]\}} \tag{18}$$

Compared to the original Eq. (2) expanded by Eq. (3)

$$R_\omega(x,t) = \rho_\omega(x,t)e^{j[2\pi\omega_0 x+\delta'(x,t)]} = \rho_\omega(x,t)e^{j\{2\pi\omega_0[x+\varphi_i(x)q_i(t)/(2\pi\omega_0)+\sum_{l=1,l\neq i}^{n}\varphi_l(x)q_l(t)/(2\pi\omega_0)]\}} \tag{19}$$

it can be seen from the right ends of Eq. (18) and Eq. (19) that the motion of the ith mode has been magnified by $(1+\alpha)$ times and the motion of other modes ($l \neq i$) have been removed, i.e., the ith mode has been blindly separated (by removing other modes in the images/video measurements) and magnified by phase manipulation as per Eq. (14) to Eq. (19). Performing such phase manipulation procedures on each subband scale ω, the reconstructed images (video)

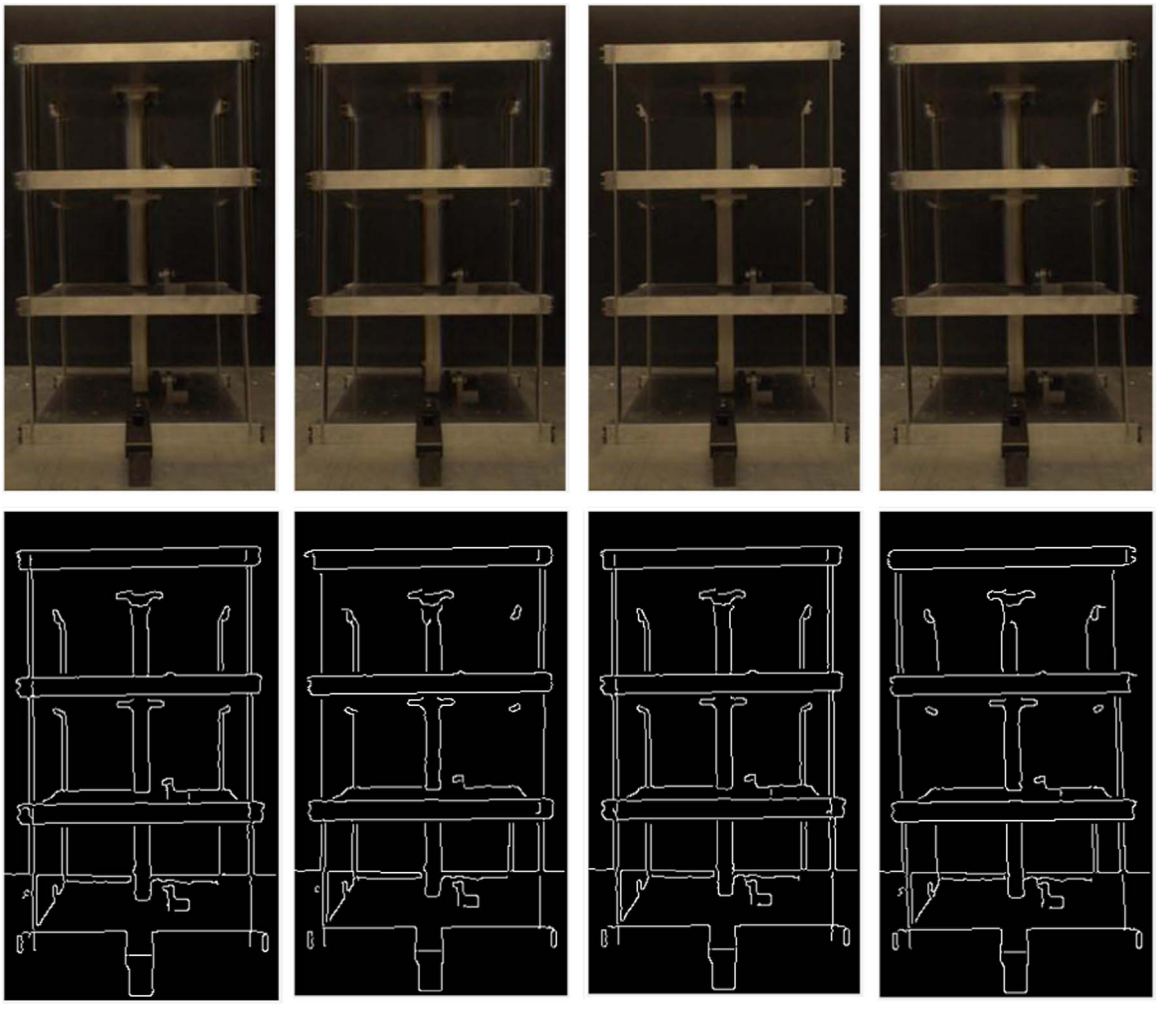

Fig. 12. A few temporal frames from the motion magnified Video (**Mode 1**, $\alpha = 8$) (**first row**, linked to **Video 2**) of the **building structure** and their corresponding mode shapes (**second row**, linked to **Video 3**) found by conducting edge detection.

$$\hat{I}\left(x+\hat{\delta}(x,t)\right)=\sum_{\omega=-\infty}^{\infty}\hat{R}_{\omega}(x,t)=\sum_{\omega=-\infty}^{\infty}\rho_{\omega}(x,t)e^{j\{2\pi\omega_0[x+(1+\alpha)\varphi_i q_i(t)/(2\pi\omega_0)]\}} \tag{20}$$

contain only the magnified vibration motion of the ith mode, whose vibration is readily visualized in $\hat{I}$. Afterwards, standard image edge detection can be performed on $\hat{I}$ to segment the structure from the background, thereby obtaining the high-resolution vibration mode shapes of the structure. Other modes can be blindly separated and magnified along the same procedures as per (Eqs. (14)–20).

2.3.2. Signal (weak mode motion) to noise ratio analysis

As per Eqs. (14)–(20), even the subtle vibration motion of the weakly-excited modes is magnified by $(1+\alpha)$ times in the reconstructed video $\hat{I}$ containing only one individual mode, which is advantageous for the following mode shape estimation by edge detection. Noise robustness of the proposed motion magnification based method is analyzed in this section to further support such a scheme.

If the image $I\left(x+\delta(x,t)\right)$ is contaminated by noise $z(x,t)$, the subband representation of the noisy image $I\left(x+\delta(x,t)\right)+z(x,t)$ becomes

$$\tilde{R}_{\omega}(x,t)=\rho_{\omega}(x,t)e^{j2\pi\omega_0\left(x+\delta(x,t)\right)}+\mathbb{N}_{\omega}(x,t) \tag{21}$$

where "$\sim$" denotes "noisy" and $\mathbb{N}_{\omega}(x,t)$ is the subband of the noise. Assuming that the amplitude of the noiseless image, $\rho_{\omega}(x,t)$, is much larger than that of $\mathbb{N}_{\omega}(x,t)$, then the extracted phase of $\tilde{R}_{\omega}(x,t)$ will still be $\psi(x,t)=2\pi\omega_0\left(x+\delta(x,t)\right)$ as well as per Eq. (3) to Eq. (17). Following Eq. (18) and applying the phase magnification to $\tilde{R}_{\omega}(x,t)$

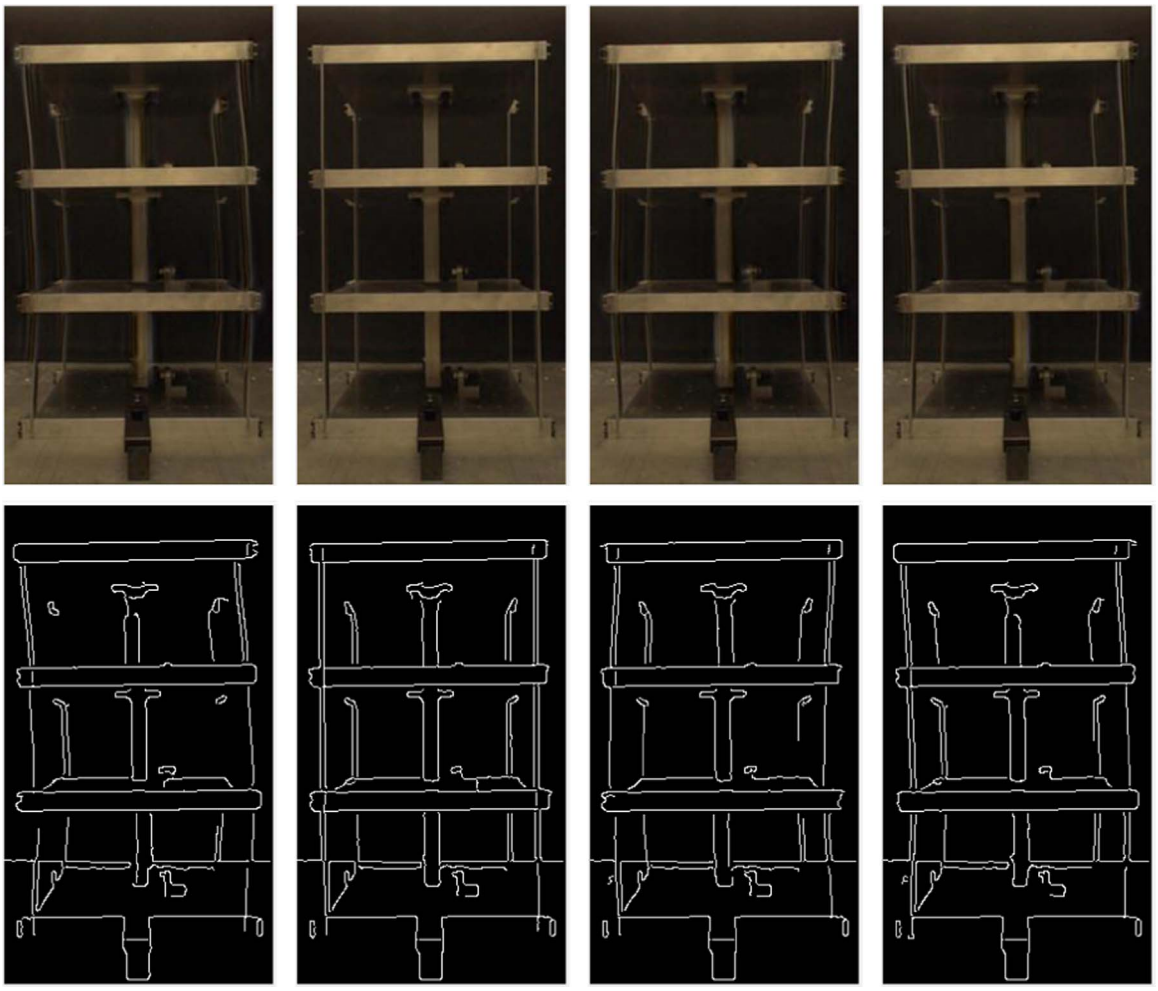

Fig. 13. A few temporal frames from the motion magnified Video (**Mode 2**, $\alpha = 40$) (**first row**, linked to **Video 4**) of the **building structure** and their corresponding mode shapes (**second row**, linked to **Video 5**) found by conducting edge detection.

$$\hat{\hat{R}}_{\omega}(x,t) = R_{\omega}(x,t)e^{j\hat{\hat{\delta}}} + \mathbb{N}_{\omega}(x,t)e^{j\hat{\hat{\delta}}} = \rho_{\omega}(x,t)e^{j\left\{2\pi\omega_0\left[x+\frac{(1+\alpha)\varphi_i q_i(t)}{2\pi\omega_0}\right]\right\}} + \mathbb{N}_{\omega}(x,t)e^{j[(1+\alpha)\varphi_i q_i(t)-2\pi\omega_0\delta(x,t)]} \tag{22}$$

it can be seen from Eq. (22) that while the vibration motion of the ith mode has been magnified by $(1+\alpha)$ times in the noiseless image, the magnification only shifts the phase of the noise component and does not change the noise amplitudes. Therefore, after collapsing the subbands, the reconstructed images (video)

$$\hat{\hat{I}} = \sum_{\omega=-\infty}^{\infty}\hat{\hat{R}}_{\omega}(x,t) = \sum_{\omega=-\infty}^{\infty}\rho_{\omega}(x,t)e^{j\left\{2\pi\omega_0\left[x+\frac{(1+\alpha)\varphi_i q_i(t)}{2\pi\omega_0}\right]\right\}} + \mathbb{N}_{\omega}(x,t)e^{j[(1+\alpha)\varphi_i q_i(t)-2\pi\omega_0\delta(x,t)]} \tag{23}$$

have $(1+\alpha)$-times magnified vibration motion of the ith mode with the same noise intensity level; this is especially advantageous for the following step of edge detection on reconstructed images (video) to extract the (magnified) mode shape of the weakly-excited mode, whose subtle vibration otherwise is easily buried in the noisy images.

2.4. Summary of the proposed method

The proposed BSS and video motion magnification based output-only modal identification method can be performed blindly and autonomously using only video measurements of the structure. Summarizing those presented as per Eqs. (1)–(20), Fig. 6 shows the flowchart which correspondingly undergoes the following steps:

(1) and (2) Perform spatial multi-scale decomposition and representation by applying complex steerable pyramid filters on each frame $I(x+\delta(x,t))$ of the video to obtain the filter response (subband representation) $R_{\omega}(x,t)$ on each spatial

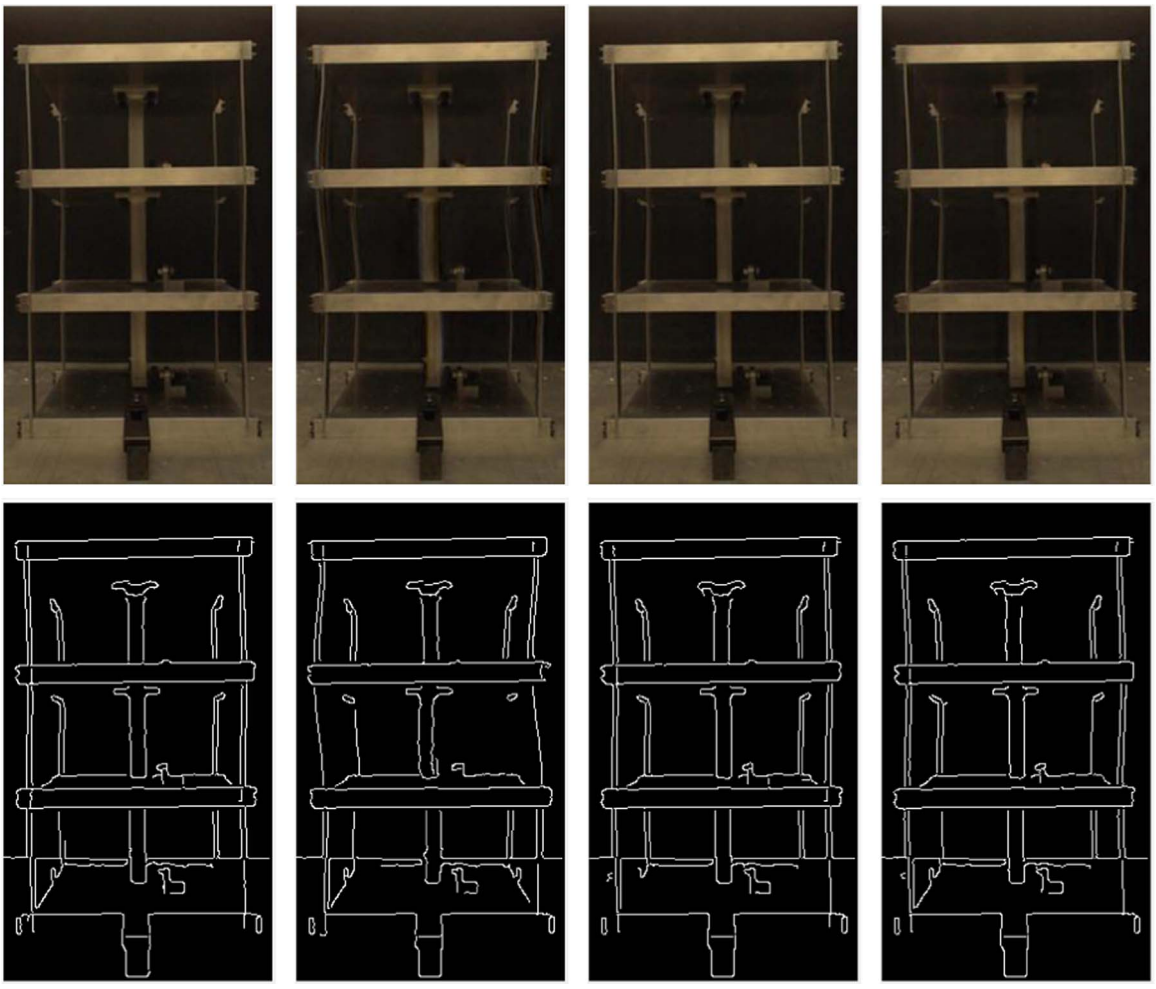

Fig. 14. A few temporal frames from the motion magnified Video (**Mode 3**, $\alpha = 70$) (**first row**, linked to **Video 6**) of the **building structure** and their corresponding mode shapes (**second row**, linked to **Video 7**) found by conducting edge detection. (Note: higher playback rate setting of the video player might be needed to correctly play Video 6 and 7.).

scale ω. From $R_\omega(x, t)$obtain the local phases $\psi(x, t)$ at each pixel of each frame. Remove the temporal mean of $\psi(x, t)$ to obtain $\delta'(x, t)$.
(3) On each scale, perform dimension reduction on $\delta'(x, t)$ using PCA to obtain r active principal components η with active non-zero singular values. Then perform BSS on η to obtain the modal coordinates $q_i(t)$ $(i = 1, \ldots, r)$, from which estimate the frequency and damping ratio.
(4) On each scale, magnifying the ith mode by multiplying $q_i(t)$ by α and $q_k(t)$ $(k \in [1, r] \neq i)$ by $\beta = -1$. Use the inverse transform from Step 3 to Step 2 to obtain the magnified $\hat{\delta}'(x, t)$ and then multiply the $R_\omega(x, t)$ by $e^{j\hat{\delta}'}$ to obtain $\hat{R}_\omega(x, t)$
(5) Reconstruct the ith-mode-magnified $\hat{I}\left(x+\hat{\delta}(x, t)\right)$by collapsing $\hat{R}_\omega(x, t)$.
(6) Perform edge detection on $\hat{I}$ to obtain the high-resolution ith mode shapes of the structure. Repeat Step 4 to Step 6 for other modes.

2.5. Distinctions with closely-related work

While the proposed method is built upon the phase-based motion magnification technique in Ref. [27], it is considerably different from the existing work. The original motion magnification proposals [27] require users to first specify the frequency band of the interested motion and magnify the content within the specified frequency band. Specifically, after obtaining the local phase $R_\omega(x, t)$as per Eqs. (1)–(2), the original phase based technique temporally filters $R_\omega(x, t)$ into a user-specified frequency band of interest to magnify the filtered component. However, the frequency band containing the motion of interest is seldom known *a priori*, requiring trials and tuning in the application [27]. Ref. [28] adopted this original proposal for modal analysis. First, displacements were explicitly computed from the obtained local phases of a few pre-defined cropped regions of the images of the structure, which were called “virtual accelerometers”. Then a traditional peak-

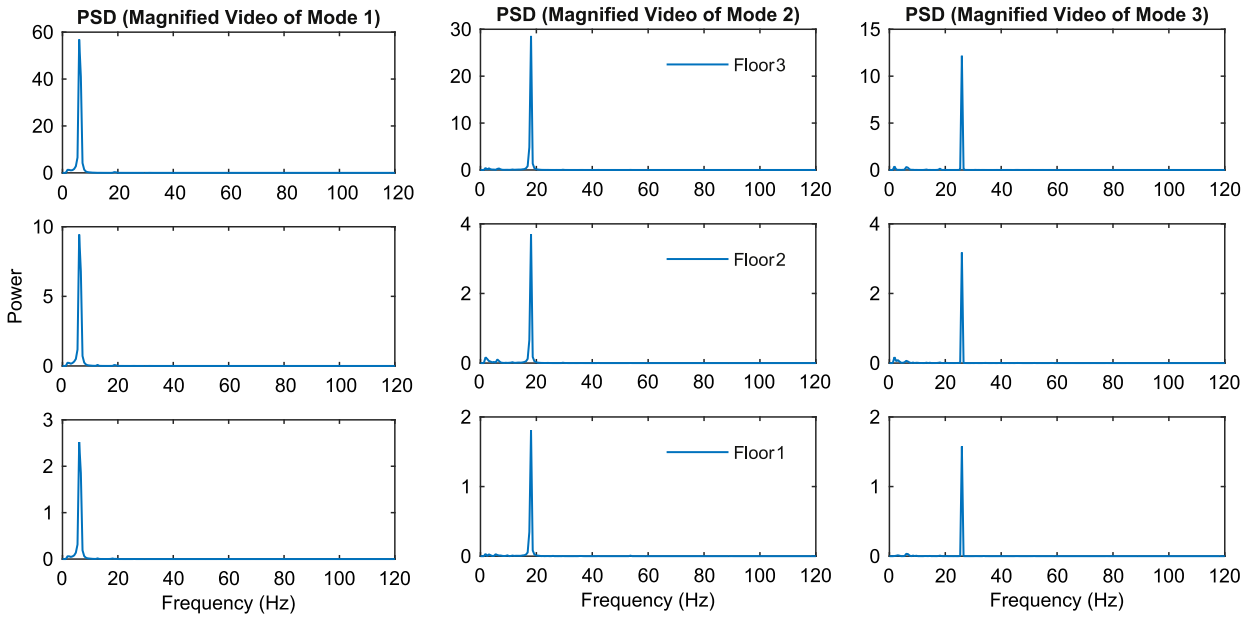

Fig. 15. The power spectral density (PSD) of the estimated motion field of some regions (distributed on different floors) of the **building structure** of the motion-magnified videos of different modes.

picking method was performed on the computed displacements to obtain the frequencies. According to the estimated frequencies, the frequency band of interest was specified and the whole process of the original phase-based motion magnification technique was then implemented. Potential limitations to such a method can be pointed out, however. First, it does not address the weakly-excited mode identification issue because the peak-picking process used in that method depends on the user's judgment of the mode, which becomes very challenging when noise is present. In addition, when closely-spaced modes are present, they will overlap in the specified frequency band, in which case that method will not be able to separate and magnify each individual mode. Also, because of the user-defined cropped regions and the peak-picking process, such procedures are not automated.

Inheriting the virtue of the unsupervised learning BSS technique, the method proposed in this study manipulates only the local phase itself and is able to blindly separate the individual modal components for modal identification and magnification with little user supervision, thus avoiding the need of user searching and specifying the frequency bands of interest (mode) and the issue of close modes identification and magnification. In addition, the proposed method does not require the explicit computation of the displacements of the structure; rather, it directly extracts all modal parameters throughout the procedure without the need of calibration or surface preparation (high-contrast markers or random speckle pattern) as compared with traditional vision based methods. Additional discussions and demonstrations are provided by the experimental examples in the following section.

3. Experimental validation

3.1. Validation on a building structure

3.1.1. Experimental setup

A bench-scale model of a three-story (discrete-type) building structure was used to validate the proposed method. The structure consists of aluminum columns and lumped mass plates on each floor, with its base fixed to a heavy solid aluminum plate on the ground. An impact hammer was used to excite the structure horizontally on the top floor. A stationary camera (Sony NXCAM with a pixel resolution of 1920×1080) mounted by a Zeiss lens with a fixed focal length of 24 mm was used to perform video measurements of the structure at a frame rate of 240 frames per second. The illumination environment is the ordinary indoor lighting condition without any external illumination enhancement. For comparisons, three uniaxial accelerometers attached on each floor were used to measure the accelerations, from which the modal parameters were also extracted and compared with those estimated from the video measurement using the proposed method. Fig. 7(a) shows the experimental setup.

3.1.2. Implementation procedures and results

For more efficient computation, the pixels were downsampled to 384×216 and a set of video measurements with less than two seconds in duration consisting of 400 frames in total (Fig. 8(a) and Video 1 in the supplementary materials) was

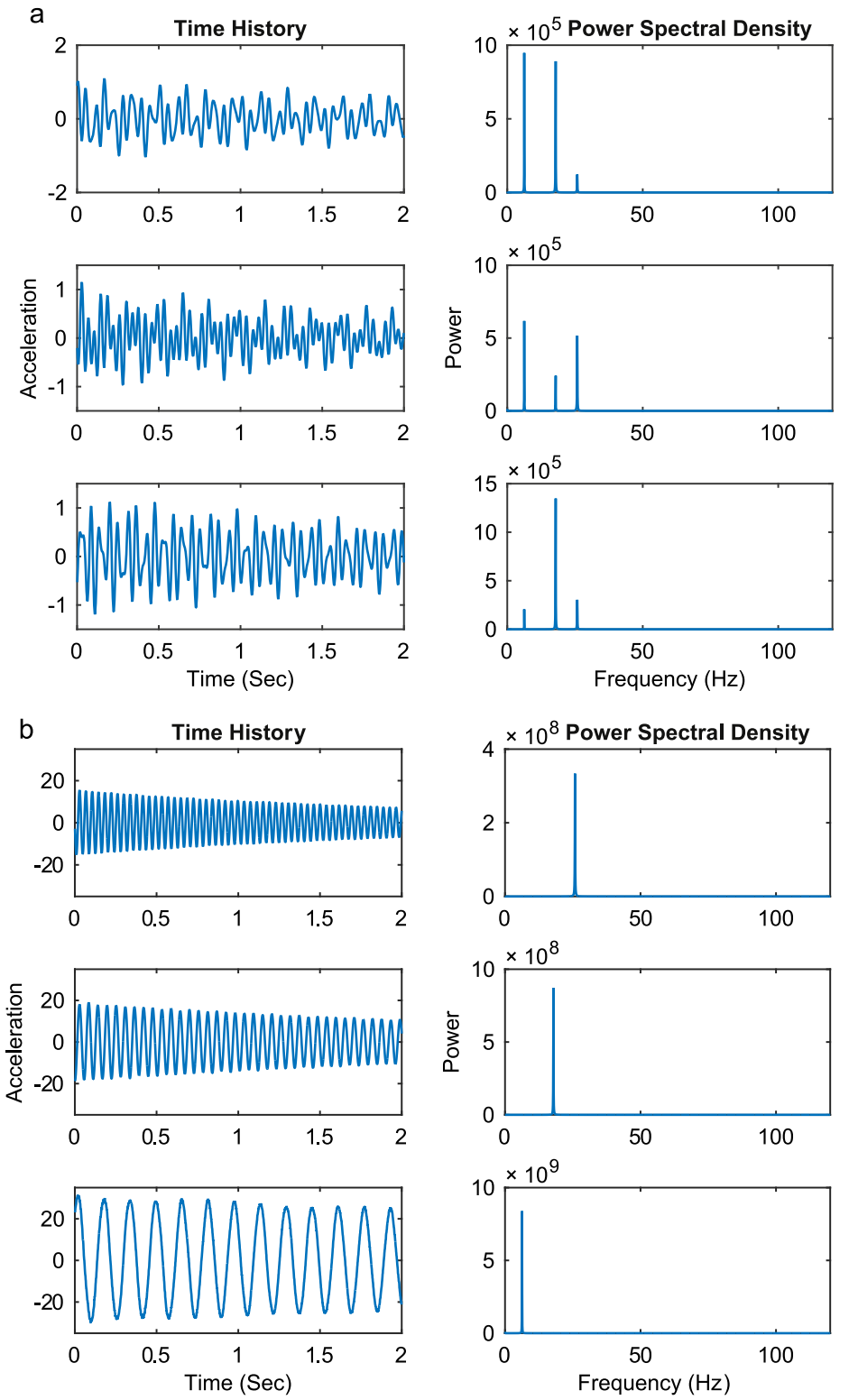

Fig. 16. (a) The measured accelerations on the three floors of the **building structure** and their power spectral density (PSD). **(b)** The estimated modal coordinates by complexity pursuit (CP) algorithm from the accelerations in **(a)** and their PSD.

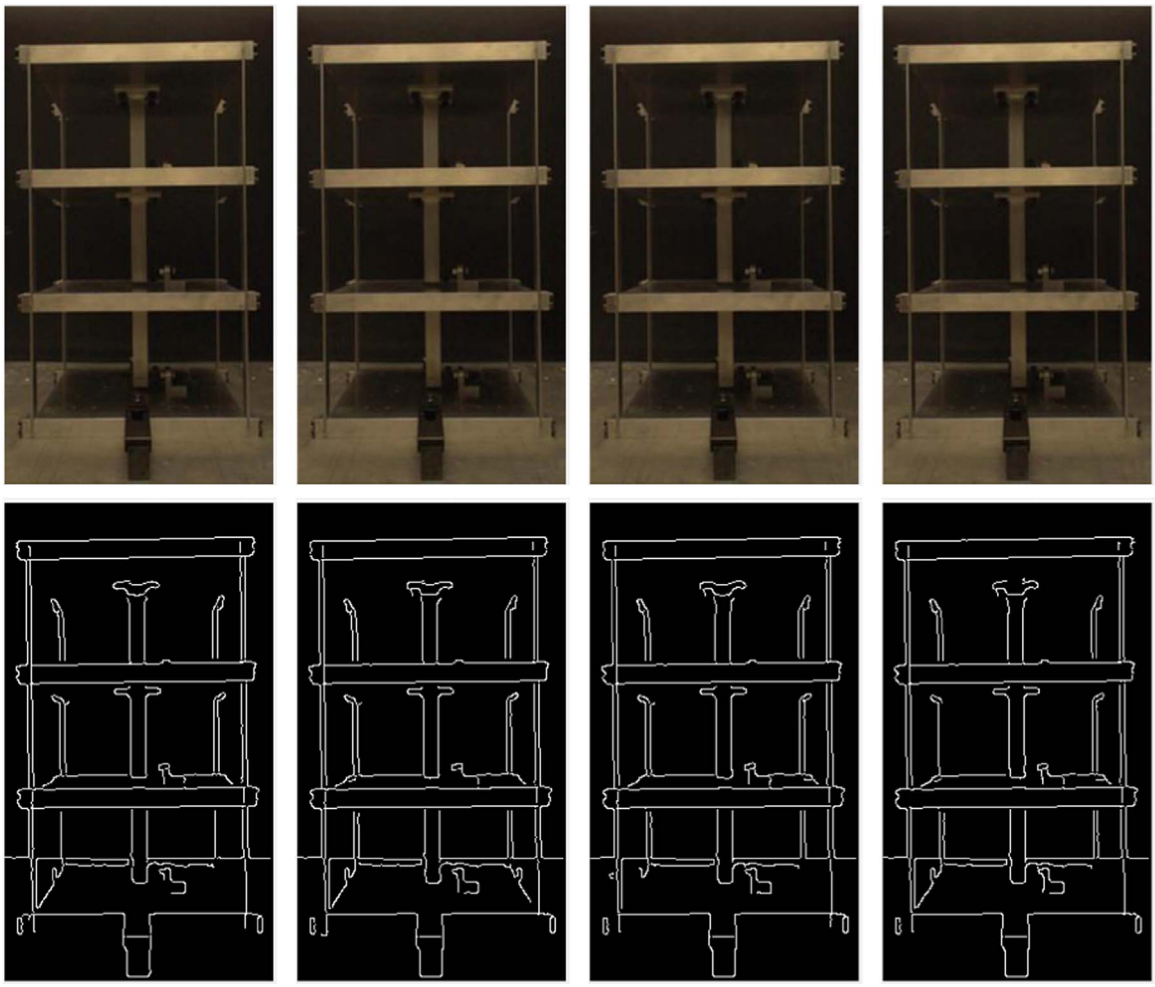

Fig. 17. A few temporal frames from the **non-magnified** Video (**Mode 3**, $\alpha = 1$) (**first row**, linked to **Video 8**) of the **building structure** and their corresponding mode shapes (**second row**, linked to **Video 9**) found by conducting edge detection. (Note: higher playback rate setting of the video player might be needed to correctly play Video 8 and 9).

used. The whole procedure of the proposed method is automated. First, each frame is decomposed and represented by a five-scale complex pyramid (subbands) with the high-pass and low-pass components. The time history of the local phases of each pixel of each frame, on each scale, were then extracted and Fig. 9 shows the time history and their power spectra of the phases of some three pixels on the first subband (Scale 1), which correspond to the structural vibration responses.

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.ymssp.2016.08.041.

On each scale, PCA is then performed on the obtained phase matrix δ'. During the PCA dimension reduction process, the eigenvalue decomposition on the covariance matrix of δ' was conducted to determine the active principal components η. For illustrations, Fig. 10 shows the eigenvalue distribution of the phase covariance matrix $\delta'\delta'^{*}$. It is seen that the eigenvalues decay radically: although the pixel dimension is on order of thousands, only the first three principal components are active, i.e., $r \approx 3$. For automated selection of the active principal components, a thresholding eigenvalue was set below 1% of the largest eigenvalue (meaning it is relatively very small) and more than 50% of the previous eigenvalue (meaning the decay has been stabilized). The three principal components $\eta(t)$ were then decomposed by the BSS technique CP, yielding the three monotone modal coordinates $q(t)$ $(i = 1,2,3)$, as shown in Fig. 11. Hilbert transform based SDOF technique and the logarithm decrement technique were applied on the estimated modal coordinates to obtain the frequency and damping ratio, respectively. These results are shown in Table 2.

On each scale, motion magnification was then applied on the estimated modal coordinates. Taking the first mode $(i = 1)$ for example: first multiply $q_1(t)$ by, say, α_1=8 and both $q_2(t)$ and $q_3(t)$ by $\beta = -1$ and then perform inverse PCA and CP transform, followed by the phase shift and reconstruction steps to yield the magnified video of Mode 1 (Video 2); the canny edge detection method was then performed on each frame of the obtained magnified video and the mode shape of Mode 1 (Video 3) was obtained. The same procedures were conducted to obtain the magnified videos and mode shapes of Mode 2 and Mode 3 (Video 4–7). For demonstrations, several frames of the video are shown in Fig. 12–14. It is seen that estimated

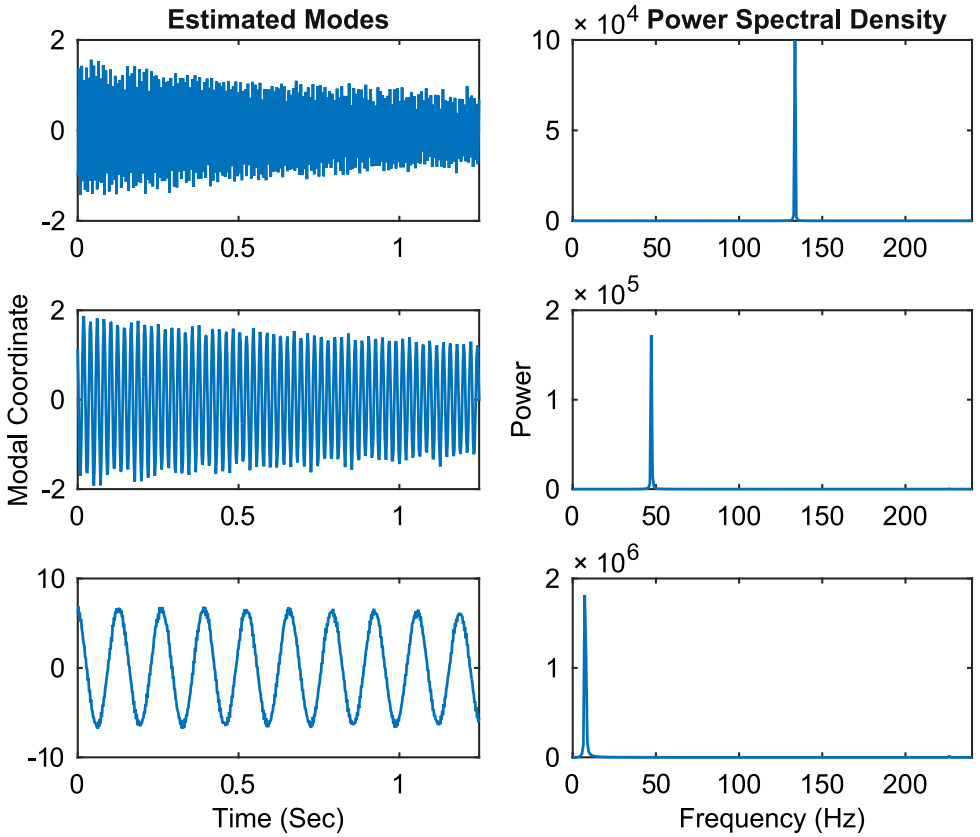

Fig. 18. The estimated modal coordinates by the complexity pursuit (CP) algorithm applied on the active principal components of the spatio-temporal pixel phases of the video of the **beam structure**.

mode shapes are consistent with the behavior (geometry) of their respective modes. To further verify the success of the separation and motion magnification, the standard Horn-Schunck optical flow method [9] was used to extract the motion field of some regions (one on each floor) of the magnified videos and their PSD's of the estimated motion are shown in Fig. 15. It is seen that each of the magnified videos indeed contains only the vibration motion of one single dominant mode, i.e., the estimated mode shape is indeed of the corresponding single mode.

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.ymssp.2016.08.041.

3.1.3. Comparisons and discussions

3.1.3.1. Comparisons with acceleration measurements. For comparisons, the measured accelerations were directly decoupled by CP into individual modal coordinates (Fig. 16), from which the frequency and damping ratio were also estimated by the Hilbert transform based SDOF method and the logarithm decrement technique. Table 2 shows that the results of the two output-only methods are in excellent agreement. As demonstrated, the proposed video measurement based method is able to estimate the very high-resolution operational mode shapes, while the accelerometer based method can only provide mode shapes at limited discrete points (three in this experiment).

3.1.3.2. Comparisons with those without magnification for weak mode identification. As discussed in Section 2.3, the phase-based motion magnification technique can significantly magnify the motion of the vibration without increasing the noise intensity level (shifting the phase of the noise instead), which is especially useful for identification of the weakly-excited mode with very subtle vibration motion that is easily contaminated by noise. In this experiment it is seen in Fig. 9 that the 3rd mode is barely present in the video measurement of the structural responses. To demonstrate the advantage of using motion magnification, the result of the 3rd mode without magnification (setting $\alpha = 1$) is shown in Fig. 17 and Video 8–9, indicating that the vibration motion is extremely weak (barely visualized) and vulnerable to noise contamination, which increase the uncertainty of the estimation accuracy of this weak mode. On contrast, evident vibration motion of the 3rd mode, otherwise very weak, is seen and can be reliably extracted from the magnified video of this mode, as shown in Fig. 14 and Video 6–7.

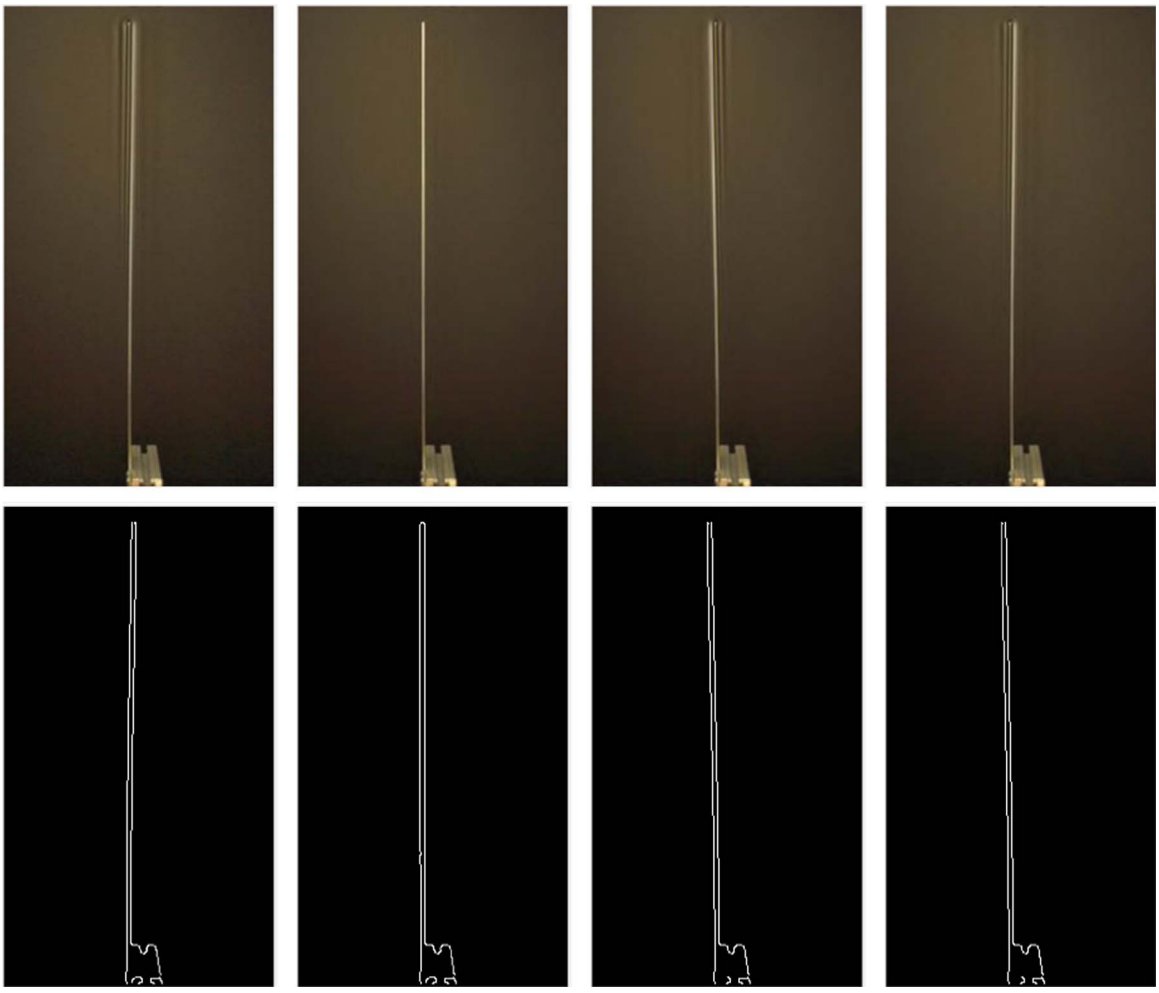

Fig. 19. A few temporal frames from the motion magnified Video of the **beam structure** (**Mode 1**, $\alpha = 8$) (**first row**, linked to **Video 11**) and their corresponding mode shapes (**second row**, linked to **Video 12**) found by conducting edge detection.

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.ymssp.2016.08.041.

3.1.3.3. Comparisons with closely-related work. Compared to the original motion magnification technique [27] and its proposal in modal analysis [28] that requires users to first specify the frequency band of the interested modal motion and magnify the component within the specified frequency band, the method proposed in this study uses BSS to blindly identify those interested modal components for magnification, thus removing the need of optical flow computation of the displacements of some cropped region that was required in Ref. [28] . In addition, the method proposed in this study directly magnifies the separated modal component with a single frequency instead of a frequency band that could incorporate multiple, close-spaced modes. Unlike the method in Ref. [28], the proposed method in this study is implemented autonomously within the modified phase-based motion magnification framework with little user inputs and supervision.

3.1.3.4. Some other issues for implementations. Most optical measurement methods such as DIC, point-tracking, and the optical flow methods are based on processing image intensity, which is known to be sensitive to illumination environment. The developed method on the other hand is based on phase manipulation, which was reported less sensitive to the light, surface, and perspective conditions [25] . Since this study mainly focuses on method development, the lab experiment in this study was conducted indoor with ordinary but stable illumination and these effects have not been particularly studied and may be subjected to further study in outdoor environment. Also in the edge detection using Canny's method which is the last step in the proposed method, distinction of the non-structure background and structure is generally required; otherwise some non-structure edges need to be removed manually or more advanced edge detection methods are needed.

In the phase-based motion magnification technique [27] , the magnified motion $(1+\alpha_i)\delta'_{(i)}$ of the ith mode is bounded by the spatial support of the transfer function (complex Gabor wavelet) on each scale. As the motion $\delta'_{(i)}$ of the weak mode (the 3rd mode in this experiment) is much smaller than others, its α_i can be set larger. In the above experimental example, α_1=8,

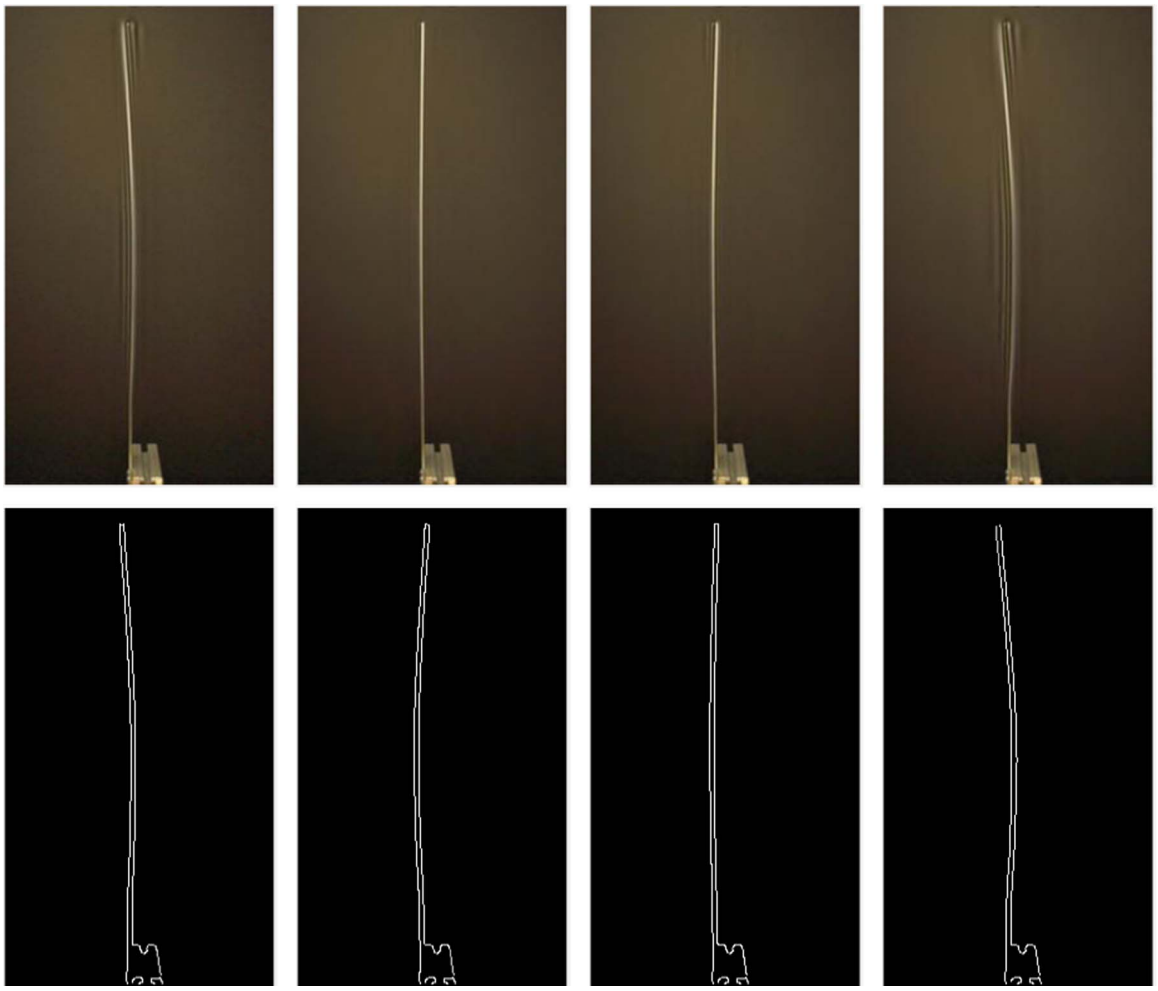

Fig. 20. A few temporal frames from the motion magnified Video of the **beam structure** (**Mode 2**, $\alpha = 25$) (**first row**, linked to **Video 13**) and their corresponding mode shapes (**second row**, linked to **Video 14**) found by conducting edge detection. (Note: higher playback rate setting of the video player might be needed to correctly play Video 13 and 14).

α_2=40, and α_3=70 were set. It was found through experimentation that the value of α can be set in a wide range (to several hundreds) with satisfactory results. Also, the steerability of the complex steerable pyramid method [29] was not utilized in this study with only one-dimensional planar vibration. Its capability of multi-orientation image analysis holds the flexibility for the method to be extended for modal analysis of multi-dimensional vibration of structures with more complex geometry.

3.2. Validation on a cantilever beam structure

A bench-scale model of an aluminum cantilever beam (continuous-type) structure was also used to validate the proposed method. An impact hammer was used to excite the beam horizontally close to the fixed end. The same camera and lens were used to perform video measurements of the beam at a frame rate of 480 frames per second (Fig. 7(b)). Similarly, the pixels were downsampled to 384×216 and a set of video measurements with less than two seconds in duration consisting of 600 frames in total (Fig. 8(b) and Video 10) was used. The procedures of the proposed method was implemented automatically. Three dominant modes were identified (Fig. 18) with modal frequencies of 7.5 Hz, 48.2 Hz, and 133.6 Hz, which are consistent with those identified from accelerations of the beam. The magnified videos of each of the three modes are demonstrated in Figs. 19–21 and Video 11–16, showing the high-resolution mode shapes. Clearly the mode shapes are consistent with those predicted by the dynamic theory of beam-type structures.

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.ymssp.2016.08.041.

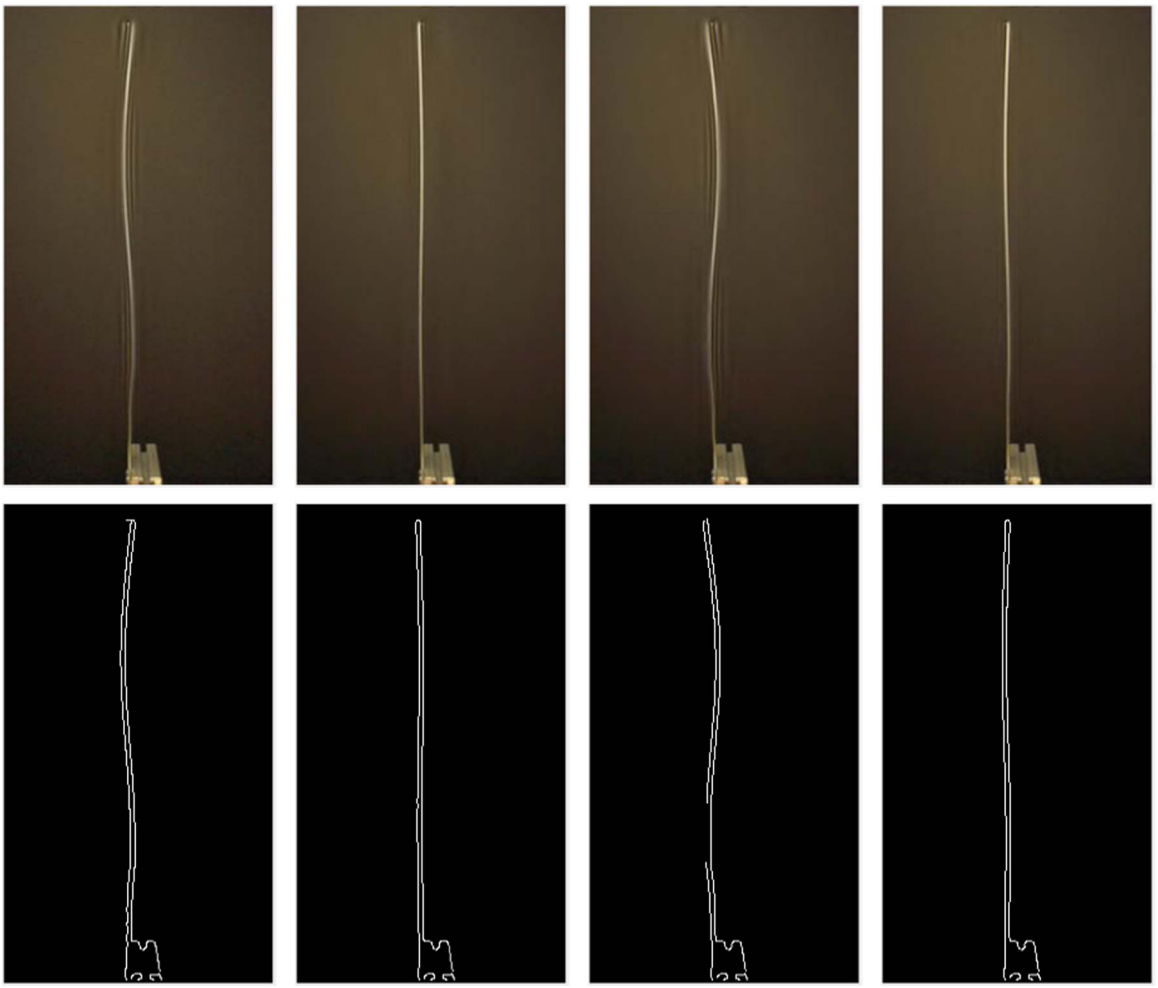

Fig. 21. A few temporal frames from the motion magnified Video of the **beam structure** (**Mode 3**, $\alpha = 45$) (**first row**, linked to **Video 15**) and their corresponding mode shapes (**second row**, linked to **Video 16**) found by conducting edge detection. (Note: higher playback rate setting of the video player might be needed to correctly play Video 15 and 16).

4. Conclusions and future work

This study develops a novel full-field, output-only (video measurement) modal analysis algorithm in the modified phase-based video motion magnification framework. It combines a multi-scale pyramid decomposition and representation method and an unsupervised learning blind source separation (BBS) technique to model and manipulate the spatiotemporal pixel phases that encode the local structural vibration in the video measurements, thus capable of blindly extracting the modal frequencies, damping ratio, and high-resolution mode shapes from line of sight video measurements of the structure. Compared to existing vision-based methods, it requires no structural surface preparation and performs in a relatively efficient and autonomous manner. Compared to the closely-related phase-based motion magnification based method, it requires minimum user input and supervision without explicit optical flow computation. Also, the noise-robustness of phase-based motion magnification is motivated and interpreted in identification and visualization of the weakly-excited mode with subtle vibration motion in the structural responses, which is a common and challenging issue in operational modal identification. Laboratory experiments are demonstrated to validate the developed method with comparisons to the accelerometer measurement based method.

While the developed method shows potential for operational modal analysis, future work needs to be conducted for its practical implementations. Real-world structures normally contain in-plane and out-of-plane vibration, which requires extending the proposed method to three-dimension. In addition, the proposed method may need to be enhanced to be robust to outdoor implementation with varying environment.

5. Acknowledgments

The authors are grateful for the support of the Los Alamos National Laboratory Lab Directed Research and Development program. This program has supported this work in the form of a Director's funded postdoctoral fellowship for Yongchao Yang and an Early Career Award for David Mascarenas. We would also like to acknowledge Jarrod McClean of the Chemistry Department at Harvard University (now at Lawrence Berkeley National Laboratory) and Nathan Sharp of the Mechanical Engineering Department at Purdue University for initiating some of the lab experiments for this study in the spring and summer of 2013 in collaboration with Los Alamos National Laboratory researchers.

References

[1] S. Doebling, C. Farrar, M. Prime, A summary review of vibration-based damage identification methods, Shock Vib. Dig. 30 (2) (1998) 91–105.
[2] W. Fan, P. Qiao, Vibration-based damage identification methods: a review and comparative study, Struct. Health Monit. 10 (1) (2011) 83–111.
[3] D. Ewins, Modal Testing: Theory, Practice and Application, Research Studies Press LTD, Baldock, Hertfordshire, England, 2000.
[4] M.I. Friswell, J.E. Mottershead, Finite Element Model Updating in Structural Dynamics, Kluwer Academic Publishers, London, 1995.
[5] A. Stanbridge, D. Ewins, Modal testing using a scanning laser Doppler vibrometer, Mech. Syst. Signal Process. 13 (2) (1999) 255–270.
[6] P. Castellini, E. Martarelli, E. Tomasini, Laser doppler vibrometry: development of advanced solutions answering to technology's needs, Mech. Syst. Signal Process. 20 (6) (2006) 1265–1285.
[7] D. Di Maio, D. Ewins, Continuous scan, a method for performing modal testing using meaningful measurement parameters: Part I, Mech. Syst. Signal Process. 25 (8) (2011) 3027–3042.
[8] M. Sutton, J. Orteu, H. Schreier, Image correlation for shape, motion and deformation measurements: basic concepts, theory and applications, Springer Science & Business Media, Berlin, Germany, 2009.
[9] B.K.P. Horn, B.G. Schunck, Determining optical flow, Artif. Intell. 17 (1981) 185–204.
[10] P. Olaszek, Investigation of the dynamic characteristic of bridge structures using a computer vision method, Measurement 25 (3) (1999) 227–236.
[11] S. Patsias, W. Staszewskiy, Damage detection using optical measurements and wavelets, Struct. Health Monit. 1 (1) (2002) 5–22.
[12] A. Wahbeh, J. Caffrey, S. Masri, A vision-based approach for the direct measurement of displacements in vibrating systems, Smart Mater. Struct. 12 (5) (2003) 785.
[13] T. Schmidt, J. Tyson, K. Galanul, Full-field dynamic displacement and strain measurement using advanced 3d image correlation photogrammetry: Part 1, Exp. Tech. 27 (3) (2003) 47–50.
[14] J. Lee, M. Shinozuka, A vision-based system for remote sensing of bridge displacement, Ndt & E Int. 39 (5) (2006) 425–431.
[15] C. Chang, Y. Ji, Flexible videogrammetric technique for three-dimensional structural vibration measurement, J. Eng. Mech. 133 (6) (2007) 656–664.
[16] J. Morlier, G. Michon, Virtual vibration measurement using KLT motion tracking algorithm, J. Dyn. Syst., Meas., Control 132 (1) (2010) 011003.
[17] E. Caetano, S. Silva, J. Bateira, A vision system for vibration monitoring of civil engineering structures, Exp. Tech. 35 (4) (2011) 74–82.
[18] M. Feng, Y. Fukuda, D. Feng, M. Mizuta, Nontarget vision sensor for remote measurement of bridge dynamic response, J. Bridge Eng. (2015) 04015023.
[19] T. Siebert, R. Wood, K. Splitthof, High speed image correlation for vibration analysis, J. Phys.: Conf. Ser. 181 (1) (2009) 012064.
[20] M.N. Helfrick, C. Niezrecki, P. Avitabile, T. Schmidt, 3D digital image correlation methods for full-field vibration measurement, Mech. Syst. Signal Process. 25 (3) (2011) 917–927.
[21] W. Wang, J. Mottershead, J. Ihle, T. Siebert, H. Schubach, Finite element model updating from full-field vibration measurement using digital image correlation, J. Sound Vib. 330 (8) (2011) 1599–1620.
[22] C. Warren, C. Niezrecki, P. Avitabile, P. Pingle, Comparison of FRF measurements and mode shapes determined using optically image based, laser, and accelerometer measurements, Mech. Syst. Signal Process. 25 (6) (2011) 2191–2202.
[23] W. Wang, J. Mottershead, T. Siebert, A. Pipino, Frequency response functions of shape features from full-field vibration measurements using digital image correlation, Mech. Syst. Signal Process. 28 (2012) 333–347.
[24] J. Baqersad, C. Niezreckir, P. Avitabile, Extracting full-field dynamic strain on a wind turbine rotor subjected to arbitrary excitations using 3D point tracking and a modal expansion technique, J. Sound Vib. 352 (2015) 16–29.
[25] D. Fleet, D. Allan, Computation of component image velocity from local phase information, Int. J. Comput. Vis. 5 (1) (1990) 77–104.
[26] T. Gautama, M. Van Hulle, A phase-based approach to the estimation of the optical flow field using spatial filtering, Neural Netw. IEEE Trans. 13 (5) (2002) 1127–1136.
[27] N. Wadhwa, M. Rubinstein, F. Durand, N. Wadhwa, Phase-based video motion processing, ACM Trans. Graph. 32 (4) .
[28] J. Chen, N. Wadhwa, Y. Cha, F. Durand, W. Freeman, O. Buyukozturk, Modal identification of simple structures with high-speed video using motion magnification, J. Sound Vib. 345 (2015) 58–71.
[29] E. Simoncelli, W.T. Freeman, The steerable pyramid: a flexible architecture for multi-scale derivative computation, in: Proceedings of the International Conference on Image Processing, Washington, DC, 1995.
[30] A. Hyvärinen, E. Oja, Independent component analysis: algorithms and applications, Neural Netw. 13 (4) (2000) 411–430.
[31] J. Portilla, E. Simoncelli, A parametric texture model based on joint statistics of complex wavelet coefficients, Int. J. Comput. Vis. 40 (1) (2000) 49–70.
[32] B. Feeny, R. Kappagantu, On the physical interpretation of proper orthogonal modes in vibration, J. Sound Vib. 211 (1998) 607–616.
[33] J. Antoni, Blind separation of vibration components: principles and demonstrations, Mech. Syst. Signal Process. 19 (2005) 1166–1180.
[34] G. Kerschen, F. Poncelet, J.-C. Golinval, Physical interpretation of independent component analysis in structural dynamics, Mech. Syst. Signal Process. 21 (2007) 1561–1575.
[35] F. Poncelet, G. Kerschen, J.-C. Golinval, D. Verhelst, Output-only modal analysis using blind source separation techniques, Mech. Syst. Signal Process. 21 (6) (2007) 2335–2358.
[36] W. Zhou, D. Chelidze, Blind source separation based vibration mode identification, Mech. Syst. Signal Process. 21 (8) (2007) 3072–3087.
[37] S. McNeill, D. Zimmerman, A framework for blind modal identification using joint approximate diagonalization, Mech. Syst. Signal Process. 22 (7) (2008) 1526–1548.
[38] A. Sadhu, B. Hazra, S. Narasimhan, Decentralized modal identification of structures using parallel factor decomposition and sparse blind source separation, Mech. Syst. Signal Process. 41 (1) (2013) 396–419.
[39] Y. Yang, S. Nagarajaiah, Blind modal identification of output-only structures in time-domain based on complexity pursuit, Earthq. Eng. Struct. Dyn. 42 (13) (2013) 1885–1905.
[40] S. Ghahari, F. Abazarsa, M. Ghannad, E. Taciroglu, Response-only modal identification of structures using strong motion data, Earthq. Eng. Struct. Dyn. 42 (8) (2013) 1221–1242.
[41] J. Antoni, S. Chauhan, A study and extension of second-order blind source separation to operational modal analysis, J. Sound Vib. 332 (4) (2013) 1079–1106.
[42] P. Brewick, A. Smyth, On the application of blind source separation for damping estimation of bridges under traffic loading, J. Sound Vib. 333 (26) (2014) 7333–7351.

[43] J. Stone, Blind source separation using temporal predictability, Neural Comput. 13 (7) (2001) 1559–1574.
[44] Y. Yang, S. Nagarajaiah, Structural damage identification via a combination of blind feature extraction and sparse representation classification, Mech. Syst. Signal Process. 45 (1) (2014) 1–23.
[45] A. Agneni, Modal parameter estimates from autocorrelation functions of highly noisy impulse responses, Int. J. Anal. Exp. Modal Anal. 7 (4) (1992) 285–297.

Yongchao Yang is a Director's Postdoctoral Fellow in the Engineering Institute at Los Alamos National Laboratory, New Mexico, United States. He obtained his Ph.D from Rice University, Houston, Texas in 2014 and bachelor's from Harbin Institute of Technology, China in 2010, both in civil engineering. His teaching and research interests are in the areas of structural dynamics, system identification, and structural health monitoring. He is particularly interested in high-resolution sensing and modeling the full-field dynamics and state-of-health of structures using computer vision and machine learning techniques. He is the author of about twenty international journal publications and three book chapters. He obtained awards of the Rice University Research SREECH Competition (2012) and the Rice University Research Nugget Competition (2012). He is the first author of the Best Paper Award of the United Nations International Conference on Sustainable Development (New York, 2015). He envisions his future research towards developing mathematical models and computational tools for improving the resilience and sustainability of urban infrastructure.

Charles "Chuck" Farrar received a Ph.D in civil engineering from the University of New Mexico in 1988. He has 30 years' experience at Los Alamos National Laboratory where he is currently the Engineering Institute Leader. His research interests focus on developing integrated hardware and software solutions to structural health monitoring problems. The results of this research are documented in more than 300 publications and numerous keynote lectures at international conferences. He teaches a graduate course in structural health monitoring at University of California, San Diego and has development of a short course entitled *Structural Health Monitoring: A Statistical Pattern Recognition Approach* that has been offered more than 20 times to industry and government agencies in Asia, Australia, Europe and the U.S. In 2003 he received the inaugural Structural Health Monitoring Lifetime Achievement Award at the International Workshop on Structural Health Monitoring and in 2012 he was elected a Los Alamos National Laboratory Fellow.

David D.L. Mascareñas earned his Ph.D and M.S in structural engineering at the University of California San Diego, in La Jolla, CA in 2008 and 2006 respectively. He received the B.S in mechanical engineering at Colorado State University in Fort Collins, CO in 2004. He worked as a laboratory manager at SAIC/Sullivan International in 2009 to develop systems health monitoring software for ground-based robots. In 2010 he was a Director's funded post-doctoral researcher at Los Alamos National Laboratory. During his time as a postdoc David focused on developing tools and techniques for ensuring the cyber-physical security of unattended robots. In 2012 he was converted to a technical staff member at the Los Alamos National Laboratory where he currently performs research on cyber-physical systems at the Engineering Institute. He currently performs research on structural health monitoring, and cyber-physical systems. In 2016 he was the recipient of a United States Presidential Early Career Award in Science and Engineering. David is particularly interested in applying structural health monitoring technologies to address sustainability issues arising from rapid urbanization and climate change. In 2015 he led the research team that wrote the paper, "Potential Structural Health Monitoring Tools to Mitigate Corruption in the Construction Industry Associated with Rapid Urbanization." This paper won the best paper award at the 2015 International Conference on Sustainable Development at Columbia University in anticipation of the 2015 UN meeting on sustainable development. David is a deputy director of the Los Alamos National Laboratory-Engineering Institute. In this role he helps run the Los Alamos Dynamic Summer School which involves mentoring undergraduate and graduate students. David is also the primary architect of the Los Alamos National Laboratory Advanced Studies Institute, a professional development program for Ph.D students and postdocs. Going forward David is particularly interested in developing computational tools to perform high-resolution infrastructure monitoring at the city scale, help augment human creativity and to further the development of his artificial personality synthesis framework.