Journal of Sound and Vibration 345 (2015) 58–71

Contents lists available at ScienceDirect

Journal of Sound and Vibration

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

Modal identification of simple structures with high-speed video using motion magnification

Justin G. Chen, Neal Wadhwa, Young-Jin Cha, Frédo Durand, William T. Freeman, Oral Buyukozturk *

Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

ARTICLE INFO

Article history:
Received 27 June 2014
Received in revised form
7 January 2015
Accepted 19 January 2015
Handling Editor: I. Lopez Arteaga
Available online 2 March 2015

ABSTRACT

Video cameras offer the unique capability of collecting high density spatial data from a distant scene of interest. They can be employed as remote monitoring or inspection sensors for structures because of their commonplace availability, simplicity, and potentially low cost. An issue is that video data is difficult to interpret into a format familiar to engineers such as displacement. A methodology called motion magnification has been developed for visualizing exaggerated versions of small displacements with an extension of the methodology to obtain the optical flow to measure displacements. In this paper, these methods are extended to modal identification in structures and the measurement of structural vibrations. Camera-based measurements of displacement are compared against laser vibrometer and accelerometer measurements for verification. The methodology is demonstrated on simple structures, a cantilever beam and a pipe, to identify and visualize the operational deflection shapes. Suggestions for applications of this methodology and challenges in real-world implementation are given.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Modal analysis of structures depends on the accurate and swift collection of data from a vibrating structure so the data can be analyzed to determine the modal characteristics. The end goal for a sensor system for modal analysis is to be able to quickly and robustly collect data from a vibrating structure. Contact accelerometers are commonly used for modal analysis and are extremely precise, however densely instrumenting a structure is difficult and tedious. Also, when the structure is small compared to the size of an accelerometer, the presence of added mass from accelerometers can affect the result. There are methods for correcting for the effects of accelerometer mass loading, but they are not exact [1,2]. Non-contact methods of measurement avoid some of these drawbacks and are being researched intensely for the purposes of modal analysis.

Non-contact methods of vibration measurement generally depend on some sort of electromagnetic radiation to transmit the information. Microwave interferometry [3] and laser methods such as laser vibrometry and electronic speckle pattern interferometry have been studied [4,5]. Cameras measuring visible light provide an interesting method for measuring vibration. They can range from precise instruments for high-frequency and high-resolution video to inexpensive units such

* Corresponding author.
E-mail address: obuyuk@mit.edu (O. Buyukozturk).

http://dx.doi.org/10.1016/j.jsv.2015.01.024
0022-460X/© 2015 Elsevier Ltd. All rights reserved.

as those on cell phones which can be chosen as necessary for the application. However, processing of videos to extract quantitative information remains relatively difficult.

Motion can be quantified in video using a number of image processing techniques. Previously researched methods use edge detection, target objects, or lights to more easily measure any structural motion [6–8]. More recent methods make use of computer vision techniques, such as the measurement of optical flow to determine the displacements of structures [9], which is related to the techniques in this paper. With a single camera, only measurements of in-plane motion can be made, however a stereo camera setup is capable of measuring both in-plane and out-of-plane motion [10]. A newer method of in-plane and out-of-plane measurement with cameras is the time-of-flight camera [11], however they currently do not provide enough resolution or speed for typical vibration measurement applications.

Recently, new computer vision techniques, collectively called motion magnification, were introduced to magnify small motions in videos [12–16]. The most recent motion magnification techniques use a signal processing approach to analyze image motions in a way analogous to an Eulerian framework for fluid flow analysis. They are ideal for computing and visualizing mode shapes because they are capable of detecting small subpixel motions that are present in the videos of vibrating structures and because they are able to separate the different modal motions through the use of temporal filtering [17]. This works well with an assumption of a stationary camera and a structure that is barely moving. The use of a video camera for vibration measurement represents a unique capability that would complement existing measurement and sensor systems currently deployed or being researched for non-destructive testing (NDT) and structural health monitoring (SHM).

The objective of this paper is to demonstrate how these newly developed motion magnification computer vision techniques can be used with high-speed camera video to visualize and quantify the vibrational mode shapes of simple structures. The theory behind motion magnification and the extraction of displacement of moving objects in videos are presented to show how videos can be analyzed qualitatively and quantitatively. An experiment comparing a high-speed camera to a laser vibrometer and accelerometer measurement was carried out to confirm that the displacements extracted from the measured video were accurate. Further experiments were conducted on a cantilever beam and a pipe with a high-speed camera. The operational deflection shapes of the cantilever beam were identified by extracting displacements from the raw recorded video and compared to those from accelerometer measurements. Additionally, the motion magnification algorithm was used to visualize the operational deflection shapes of the cantilever beam and pipe test specimens, and screenshots of the generated videos are shown and discussed. A technique for deriving the operational deflection shapes directly from the motion magnified video without extracting displacement through the use of an edge detection algorithm is also demonstrated and compared to accelerometer measurements.

2. Theory and methods

Videos are made up of a series of images and thus have two domains: the spatial domain corresponding to a 2D field of brightness values in a single image, and the time or temporal domain corresponding to the image as it evolves in time to make a video. Images can be decomposed in the spatial domain by filters into amplitude and phase signals, similar to how an accelerometer signal can be decomposed by a Fourier or wavelet transform. These spatial amplitude and phase signals from all the images of the video can be combined to obtain time-varying signals representative of the video that can be more easily manipulated than raw pixel values.

2.1. Motion magnification

Motion magnification is an algorithm for the amplification of small motions in videos. The basic principle behind the algorithm is to obtain a representation for the video such that signals representing the motions of objects in the video can be time–frequency bandpass filtered, amplified, and reconstructed back into a video where the apparent motion of objects in the video is larger in a certain frequency band. The phase-based motion magnification algorithm [14] decomposes the signal of a video into the local spatial amplitude and phase using a complex-valued steerable pyramid filter bank [18]. The local spatial phase signals are temporally Fourier decomposed into a series of sinusoids representing harmonic motion. The phase signals are then temporally bandpass filtered, amplified, and recombined back to form a motion magnified video. The result is that the video then has magnification of the motion in a specified band of temporal frequencies. A diagram of this processing workflow is shown in Fig. 1 and a detailed explanation and derivation of the procedure is given in reference [14]. For a vibrating structure, if the band of frequencies is appropriately chosen, the result will be a video of the structure's operational deflection shape at one of its approximate resonant frequencies, assuming that the modes are sufficiently spaced. In the case of closely coupled modes, the mode with a larger amplitude will likely visually dominate. An in-depth discussion of the differences and interpretation of operational deflection shapes as opposed to classical mode shapes is given in reference [19]. This qualitative visualization of a structure's operational vibration modes can be invaluable in the case of more geometrically complex structures.

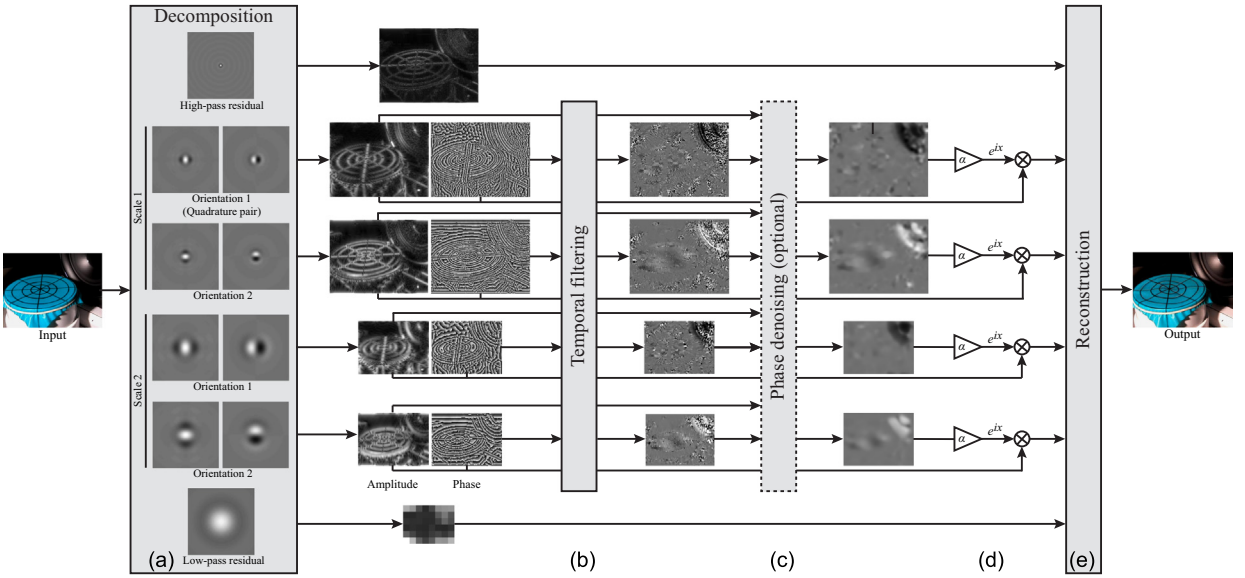

Fig. 1. Motion magnification processing workflow: (a) complex steerable pyramid filters decompose the video into amplitude and phase at different scales, (b) the decomposed phases are bandpass filtered in frequency, (c) amplitude-weighted smoothing is applied, (d) the bandpassed phases are amplified or attenuated, and (e) the video is reconstructed [14].

2.2. Operational deflection shapes directly from motion magnified video

The operational deflection shapes can be determined directly from the phase-based motion magnified video. The motion magnified video is paused during a frame with maximum deflection of the structure and a Canny edge detection routine is implemented to determine the edges of the structure. The results of the edge detection routine are cleaned up by hand because the routine leaves many artifacts that are not representative of the operational deflection shape and the video also may include other errant objects such as wires or shadows that are not of interest. This process in its current form is not expected to produce accurate results or be easily automated, but provides a quantitative measure that can be used for objective analysis and comparison to the operational deflection shapes derived from accelerometer data. For more accurate results, the displacement time series of a vibrating structure can be extracted from video using a related method described in the next section.

2.3. Displacement extraction

Using a technique related to phase-based motion magnification, from a video of a vibrating structure, the displacement signal can be computed everywhere on the structure [14]. The displacement signal is well defined only at edges in the video and then only in the direction perpendicular to the edges. This is because the motion of textureless, homogeneous regions is locally ambiguous, and determining the motion in textureless regions is an open problem in computer vision known as dense optical flow [20,21]. For the purposes of modal identification of simple structures, we can derive its motion from the motion at the edges. In the case of the cantilever beam, the entire beam is an edge and the displacement signal can be determined everywhere on it. For more complex structures with irregular geometry or coupled modes, a stereo camera setup or a textured surface is likely necessary to obtain more displacement information for accurate modal detection. We use a technique based on local phase and local amplitude in oriented complex spatial bandpass filters to simultaneously compute the displacement signal and edge strength [22,23]. The edge strength can then be used to perform a spatially local weighted average of the displacement signal to improve the signal-to-noise ratio (SNR) as in phase-based motion magnification [14].

The spatial local phase and local amplitude are quantities analogous to the phase and amplitude of Fourier series coefficients. The phase controls the location of basis function while the amplitude controls its strength. In the case of the Fourier transform, the phase corresponds to global motion. Local phase gives a way to compute local motion [22]. For a video, with image brightness specified by $I(x,y,t)$ at spatial location (x,y) and time t, the local phase and local amplitude in orientation θ at a frame at time t_0 is computed by spatially bandpassing the frame with a complex filter $G_2^\theta + iH_2^\theta$ to get

$$A_\theta(x,y,t_0)e^{i\phi_\theta(x,y,t_0)} = (G_2^\theta + iH_2^\theta) \otimes I(x,y,t_0) \quad (1)$$

where $A_\theta(x,y,t_0)$ is the local amplitude and $\phi_\theta(x,y,t_0)$ is the local phase. The filters G_2^θ and H_2^θ, shown in Fig. 2 and specified explicitly in Appendix A, are convolution kernels for processing the video frame and represent a quadrature pair that differs in phase by 90° [24]. These filters are used rather than the steerable pyramid filter bank used in motion magnification to

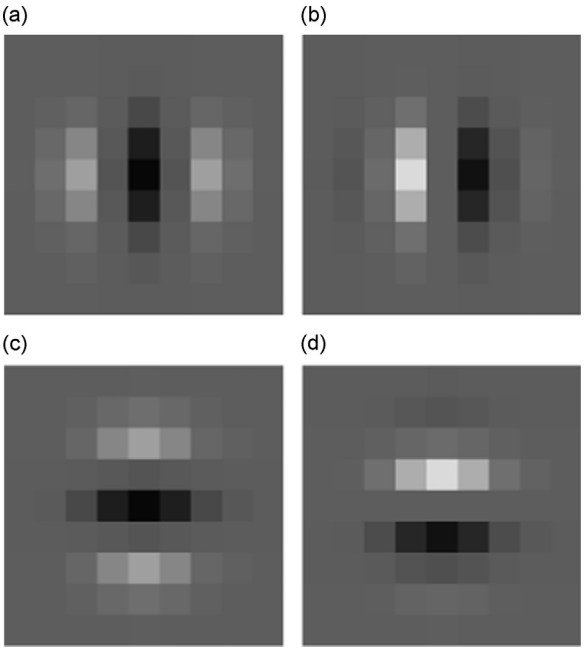

Fig. 2. Filters used to compute local phase and local amplitude: (a) real horizontal (G_2^0), (b) imaginary horizontal (H_2^0), (c) real vertical ($G_2^{\pi/2}$), and (d) imaginary vertical ($H_2^{\pi/2}$). These images represent a 9 by 9 grid of numbers where the gray level corresponds to the value of the filter, given explicitly in Appendix A.

reduce computation time, as the motion is assumed to be captured using a single spatial scale. In practice, any quadrature pair of filters could be used to determine the local amplitude and local phase, different pyramid levels could be used to capture the signals at various spatial scales, and steerability is useful for motion in more than one direction [18,24]. To increase SNR and change the scale on which the filters are operating, the video sequence is downsampled four times in each dimension spatially prior to application of the filters.

It has been demonstrated that constant contours of the local phase through time correspond to the displacement signal [22,23]. Using the notation of Eq. (1), this can be expressed as

$$\phi_\theta(x,y,t) = c \tag{2}$$

for some constant c. Differentiating with respect to time yields

$$\left(\frac{\partial\phi_\theta(x,y,t)}{\partial x}, \frac{\partial\phi_\theta(x,y,t)}{\partial y}, \frac{\partial\phi_\theta(x,y,t)}{\partial t}\right) \cdot (u,v,1) = 0 \tag{3}$$

where u and v are the velocity in the x and y directions respectively. It is approximately the case that $\partial\phi_0(x,y,t)/\partial y \approx 0$ and $\partial\phi_{\pi/2}(x,y,t)/\partial x \approx 0$. The velocities in units of pixel are

$$u = -\left(\frac{\partial\phi_0(x,y,t)}{\partial x}\right)^{-1} \frac{\partial\phi_0(x,y,t)}{\partial t} \tag{4}$$

and

$$v = -\left(\frac{\partial\phi_{\pi/2}(x,y,t)}{\partial y}\right)^{-1} \frac{\partial\phi_{\pi/2}(x,y,t)}{\partial t} \tag{5}$$

The velocity between the i th frame and the first frame for all i is computed to give a displacement signal in time. The SNR of this signal is increased by performing a spatially local weighted average of the displacement signal using the local amplitude as weights. The displacement signal is converted to units of millimeters by multiplying by the length of an object in the scene divided by the number of pixels it spans. This conversion between pixel and millimeter displacement depends on the depth of the object in the scene, and is constant for objects at the same depth, assuming no large lens distortions. In summary, each individual image of the video is processed with spatial filters to obtain spatially local phase signals sequenced in time that can represent a displacement signal of a moving object in the video. The result of the aforementioned processing is a displacement signal at all points in the image.

2.4. Video processing procedure

To analyze videos from a vibrating structure the processing methods are used as follows. The displacement extraction algorithm can be independently used on various cropped regions of the video to extract the displacement of the structure in those regions. Obtaining a displacement signal for these cropped regions gives an improvement in the signal SNR because of the spatially local weighted average, and the cropped regions can be thought of as "virtual accelerometers" that are placed on the structure that give the displacement signal for that location on the structure. These displacement signals can be analyzed in the typical fashion as one would with accelerometer data. The displacements are then analyzed to obtain resonant frequencies of the structure and operational deflection shapes are generated from these displacement signals by taking the normalized values of the frequency peaks and phase of the cross spectrum with a specified reference displacement. The resonant frequencies of the extracted displacement at the well chosen reference location are also used to determine the specific frequency bands to motion magnify the video for visualizing the operational deflection shapes. Phase-based motion magnification is applied at these frequency bands with amplification factors approximately inversely proportional to the relative amplitudes of the resonant frequency peaks to generate videos of the operational deflection shapes with sufficient movement for visual interpretation [14]. All steps of this process are currently automated except for the following automatable steps of picking the resonant frequency peaks and clean up of the mask that determines the locations of valid displacements.

2.5. Experimental method

The experiments carried out include a verification test, a cantilever beam measurement, and a pipe measurement. All three experiments were carried out in a laboratory under controlled conditions. Objects were filmed with a Phantom v10 high speed camera. Extra lighting was used so that the high-speed camera would be able to see the vibrating structures and a white poster board was typically used as a backdrop for a clutterless background for contrast. Resolutions for the video were chosen to give as much spatial data as possible and frame rates were chosen to be sufficient for the presumed vibration frequencies of the structures while still leaving enough exposure time so that the camera collects enough light. The camera was positioned such that the structure is measured dead-on to avoid angle effects and from a far enough distance to avoid any significant lens distortion. In general the experiments were controlled so that the methodology could be demonstrated in ideal conditions as a first step towards real world measurements that involve possibly unfavorable conditions.

3. Verification test

3.1. Experimental setup

In order to validate the motion magnification algorithm applied to videos of vibrating structures for the measurement of displacements and operational deflection shapes, an experiment was formulated to compare the results to standard sensors. An accelerometer was mounted on the free end of a steel cantilever beam, and the motion of the accelerometer was simultaneously measured by the accelerometer itself, a laser vibrometer, and a high speed camera, as shown in Fig. 3(a). A screenshot of the video from the camera is shown in Fig. 3(b), and the resolution of the camera was 480×288, and the

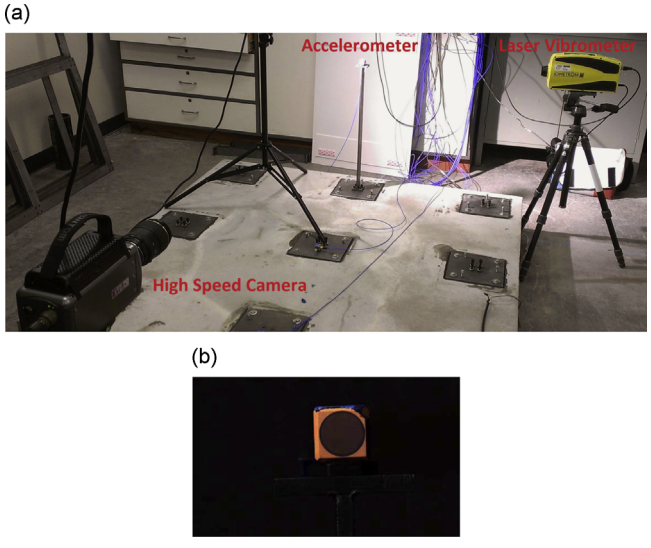

Fig. 3. Verification test: (a) experimental setup and (b) scene from high speed camera video (artificially brightened in text for visibility).

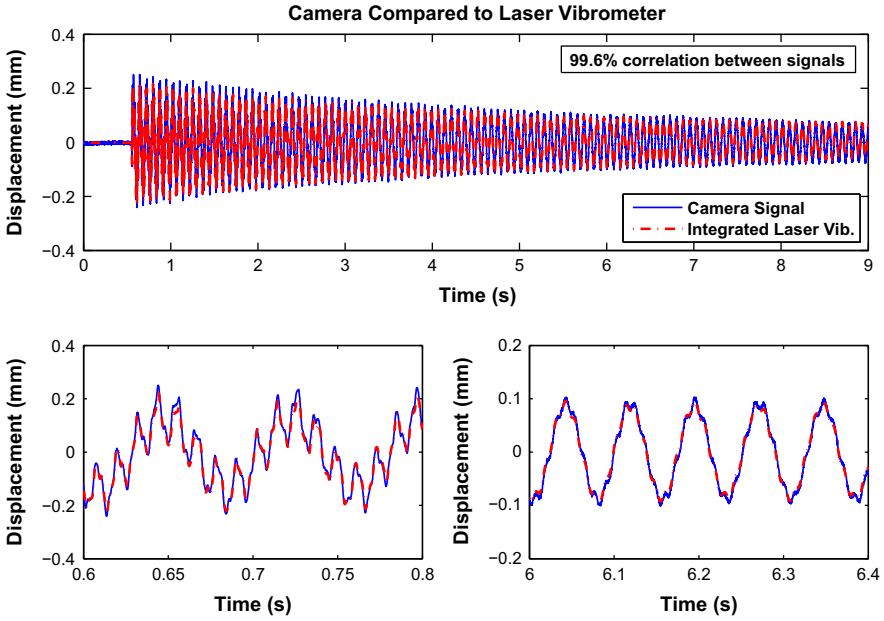

Fig. 4. Comparison between displacements derived from the camera and the laser vibrometer.

frame rate was 5000 frames per second. In the plane of the accelerometer, the video frame was 104 mm wide and 62 mm tall, with 4.615 pixels per millimeter. The cantilever beam was excited with an impact hammer near the base, and the subsequent vibration was measured for comparison. The velocity time series from the laser vibrometer was integrated to obtain displacement to verify against the displacements derived from camera measurements of the optical flow of the accelerometer movement. Time synchronization was not possible between the camera and laser vibrometer data set, so in the data the time series were aligned by eye. A fast Fourier transform (FFT) was performed on the data from the laser vibrometer, accelerometer, and camera derived displacement, and integrated if necessary to obtain the displacement so that the frequency peaks and noise floors could be directly compared.

3.2. Results

The data collected from the verification measurement were the various signals of the accelerometer movement as mounted on the cantilever beam as measured by a laser vibrometer, an accelerometer, and a high speed camera. The displacement from the camera was extracted using the local phase of the measured video as described in the theory and derivation section. The raw signals collected were the displacement time series from the camera, velocity time series from the laser vibrometer, and acceleration time series from the accelerometer. To directly compare the time series results, the laser vibrometer velocity time series was integrated numerically and results are shown in Fig. 4. The double numerical integration of the acceleration signal in the time domain to obtain displacement did not give satisfactory results due to low-frequency contamination, and is therefore not shown. The issue of low-frequency contamination in a double integrated accelerometer signal is a known issue and motivates direct measurement of displacements when possible [25]. The displacement derived from the camera closely matches the integrated laser vibrometer displacement for the whole 9 s of data, however the camera signal does show less detail and has more noise than the laser vibrometer signal. For a quantitative comparison the correlation between the camera and laser vibrometer displacement signals is 99.6 percent.

The signals were also compared in the frequency domain by using the FFT and then integrating the accelerometer and laser vibrometer signals in the frequency domain to obtain the displacements for each of the measurement methods. This allowed for easier comparison of the noise floor of the displacement derived from the camera with the other measurement methods. The comparison plot is shown in Fig. 5. Both the accelerometer and laser vibrometer data show eight resonant frequencies above the noise floor from 0 to 2500 Hz, while the camera only shows the first four resonant frequencies of the cantilever beam. The noise floor of the camera for this 9 s measurement is 40 nm, while the laser vibrometer has a noise floor of 0.2 nm, and the accelerometer has a noise floor of 0.02 nm. Given the conversion factor of 480 pixels for 104 mm, and accounting for the length of the measurement, the noise floor of the camera for this measurement was approximately 1×10^{-5} pixels per root Hertz. Possible methods for improving the noise floor of the camera measurement are to increase the number of pixels per millimeter in the frame or to capture more spatial scales with different levels or spatial downsampling of the video. Both increasing the number of pixels and capturing more spatial scales require more video processing time. Increasing the number of pixels may also reduce the achievable frame rate as there is a limit due to the

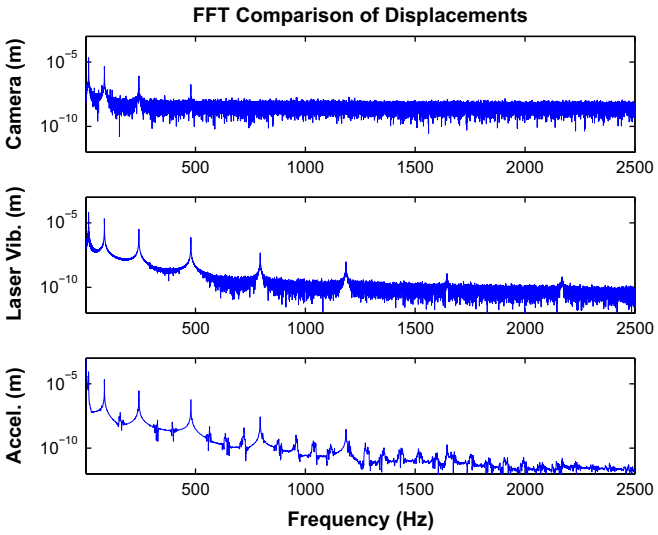

Fig. 5. Frequency space comparison between displacements derived from the camera, laser vibrometer, and accelerometers.

camera hardware. This verification measurement gives us reasonable confidence that the displacement extracted from video is accurate and that the operational deflection shapes of a column can be successfully measured by a high-speed camera.

4. Cantilever beam test

4.1. Experimental setup

To test the camera as a sensor for determining the operational deflection shapes of structures, an experiment measuring a steel cantilever beam was formulated. The cantilever beam was instrumented with nine accelerometers so that the operational deflection shapes calculated from the accelerometer data could be compared with those extracted from the high-speed camera video. The beam was excited with an impact hammer near the base and the subsequent vibration was measured by the camera and accelerometers. Fig. 6(a) shows a side view of the instrumented cantilever beam, and Fig. 6(b) shows a screenshot from the camera video. Video was taken with the camera at 1500 frames per second, and the resolution of the video was 200×1056. In the plane of the column, the video frame was approximately 124 mm wide and 655 mm tall, with 1.613 pixels per millimeter.

4.2. Operational deflection shapes from displacement extraction

The method for identifying the operational deflection shapes of the cantilever beam is analogous to the method used to extract displacement from the high speed camera video in the verification test. The accelerometer instrumented cantilever beam was measured with the high speed camera and the video was cropped to separate regions containing only one of each of the nine accelerometers. A displacement signal was independently determined from the videos for each of these regions giving an equivalent video-derived displacement signal for each accelerometer location. These displacement signals are then processed in a typical fashion, as previously described, to determine the operational deflection shapes. The camera was capable of discerning the vibration of the first four resonant frequencies of the column at 12.5 Hz, 80 Hz, 226 Hz, and 444 Hz, and thus these four operational deflection shapes were extracted. The operational deflection shapes extracted from the camera measurement are shown in Fig. 7(a). The cantilever beam was also instrumented with accelerometers to determine the operational deflection shapes from a conventional method of measurement for comparison. The operational deflection shapes calculated from the accelerometer data are shown in Fig. 7(b). The camera extracted operational deflection shapes compare well to those calculated from the accelerometer data, and both sets of extracted operational deflection shapes match the expected theoretical mode shapes of a cantilever beam. The MAC values for the comparison between the operational deflection shapes from the camera data and accelerometer data are shown in Table 1.

4.3. Visualization of operational deflection shapes with motion magnification

The phase-based motion magnification algorithm was applied to the source video with frequency bands of 12–13 Hz, 78–82 Hz, 222–230 Hz, and 438–450 Hz and amplification factors of 20, 50, 100, and 300 [14]. Screenshots of the motion magnified videos are shown in Fig. 8. These visualized operational deflection shapes also match the expected cantilever beam mode shapes. The results from the edge detection routine to determine a quantified operational deflection shape from

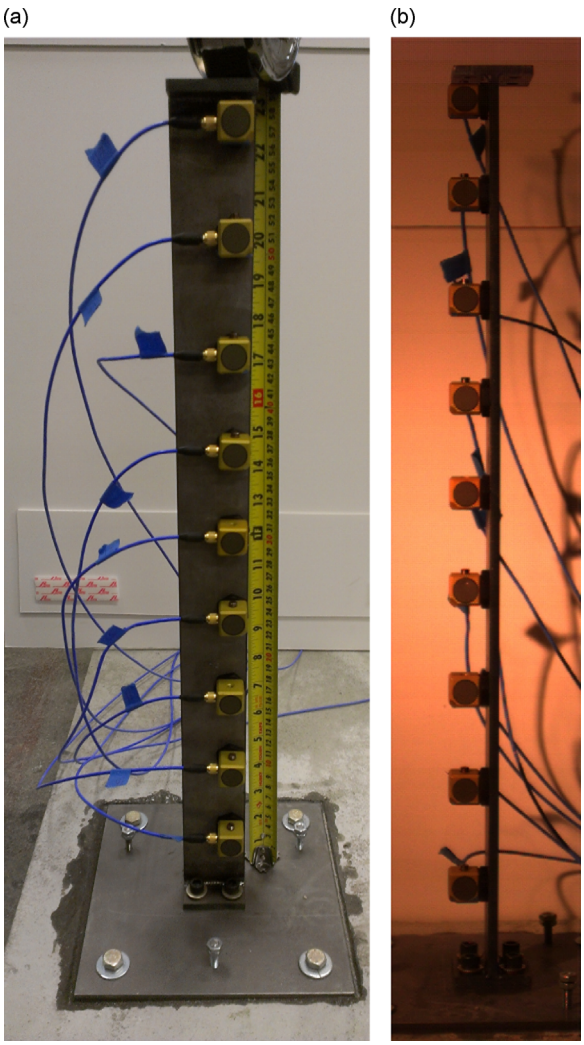

Fig. 6. Cantilever beam experimental setup: (a) instrumented cantilever beam side view and (b) screenshot from video camera.

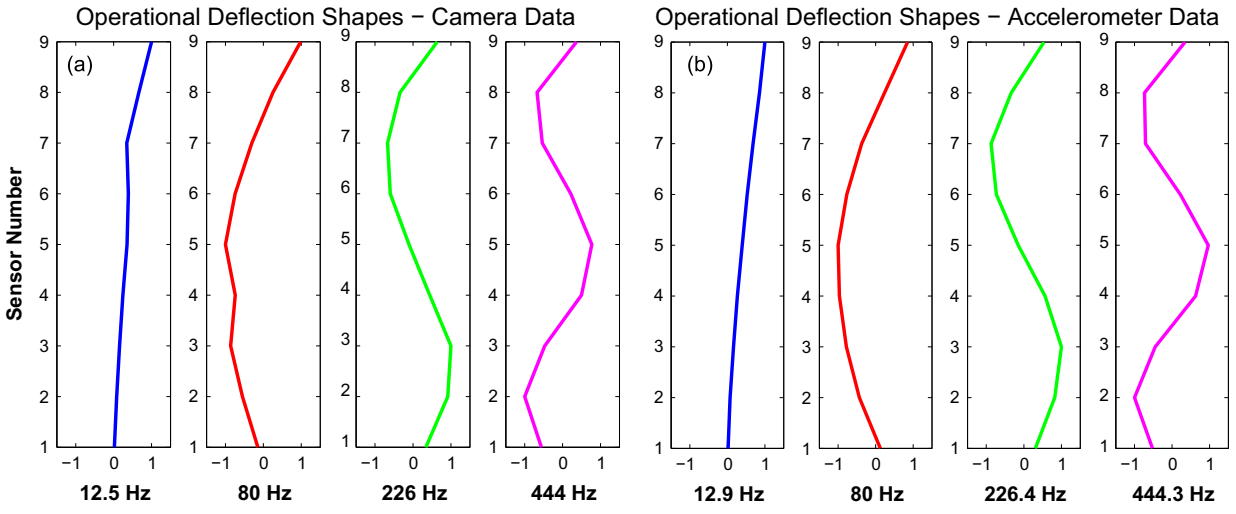

Fig. 7. Normalized operational deflection shapes of the cantilever beam from (a) displacements extracted from camera video and (b) accelerometer data.

Table 1
MAC values for the comparison between the accelerometer and camera displacement extraction derived operational mode shapes.

Accel (Hz)	Camera			
	12.5 Hz (%)	80 Hz (%)	226 Hz (%)	444 Hz (%)
12.5	95.58	0.28	0.13	0.36
80	0.34	97.78	0.91	0.25
226	0.02	0.88	97.56	2.54
444	0.02	0.25	2.14	98.46

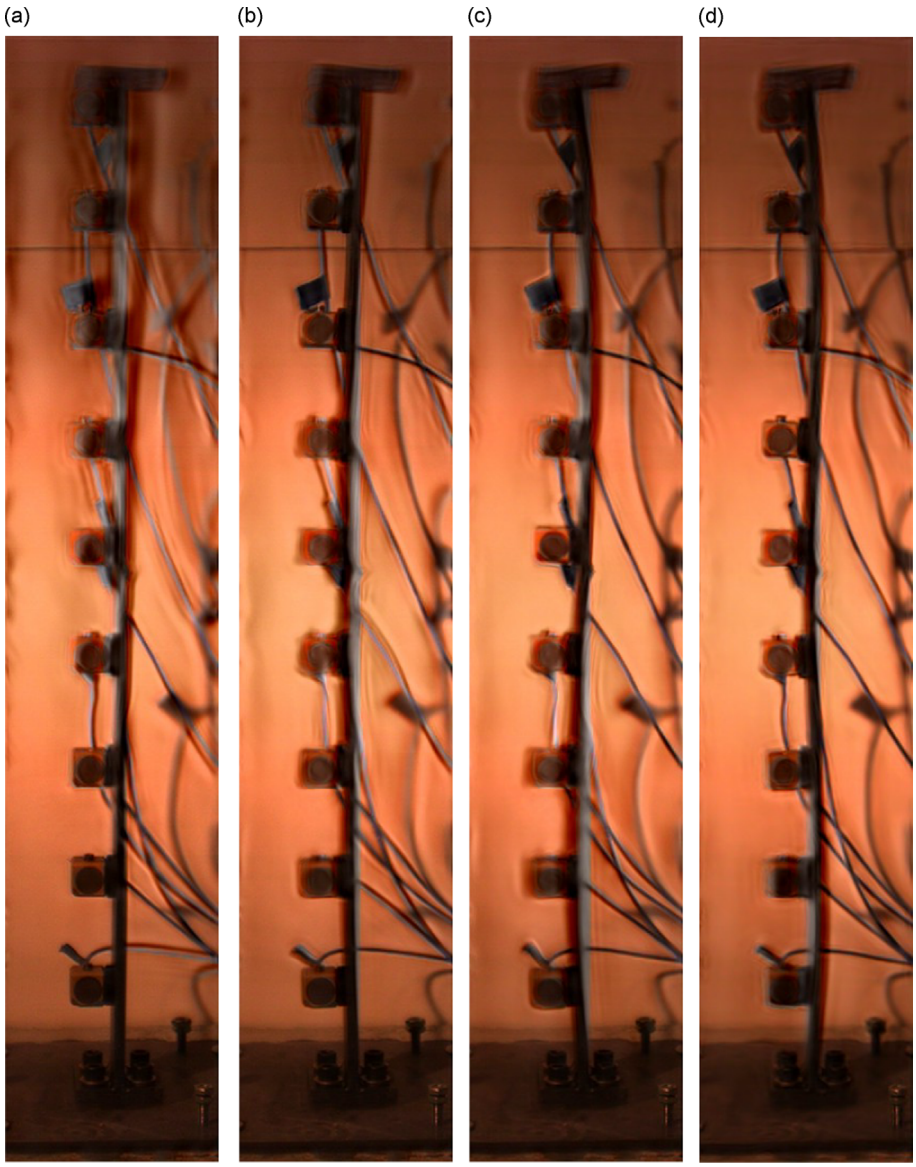

Fig. 8. Screenshots from motion magnified video of the cantilever beam: (a) 12.5 Hz, (b) 80 Hz, (c) 226 Hz, and (d) 444 Hz.

the screenshots are shown in Fig. 9 and the MAC values as compared to the accelerometer data are given in Table 2. This represents a slightly different method of deriving operational deflection shapes from video, and gives a quantitative comparison of the shapes compared to the ones extracted from accelerometer data. They do not compare as well as the operational deflection shapes derived from the video extracted displacements, however this may be due to problems with the edge detection algorithm and the clutter of the accelerometers and cables in the video. For a better visualization of the motion magnified video, a series of screenshots from the video in the frequency band of 222–230 Hz are shown in Fig. 10.

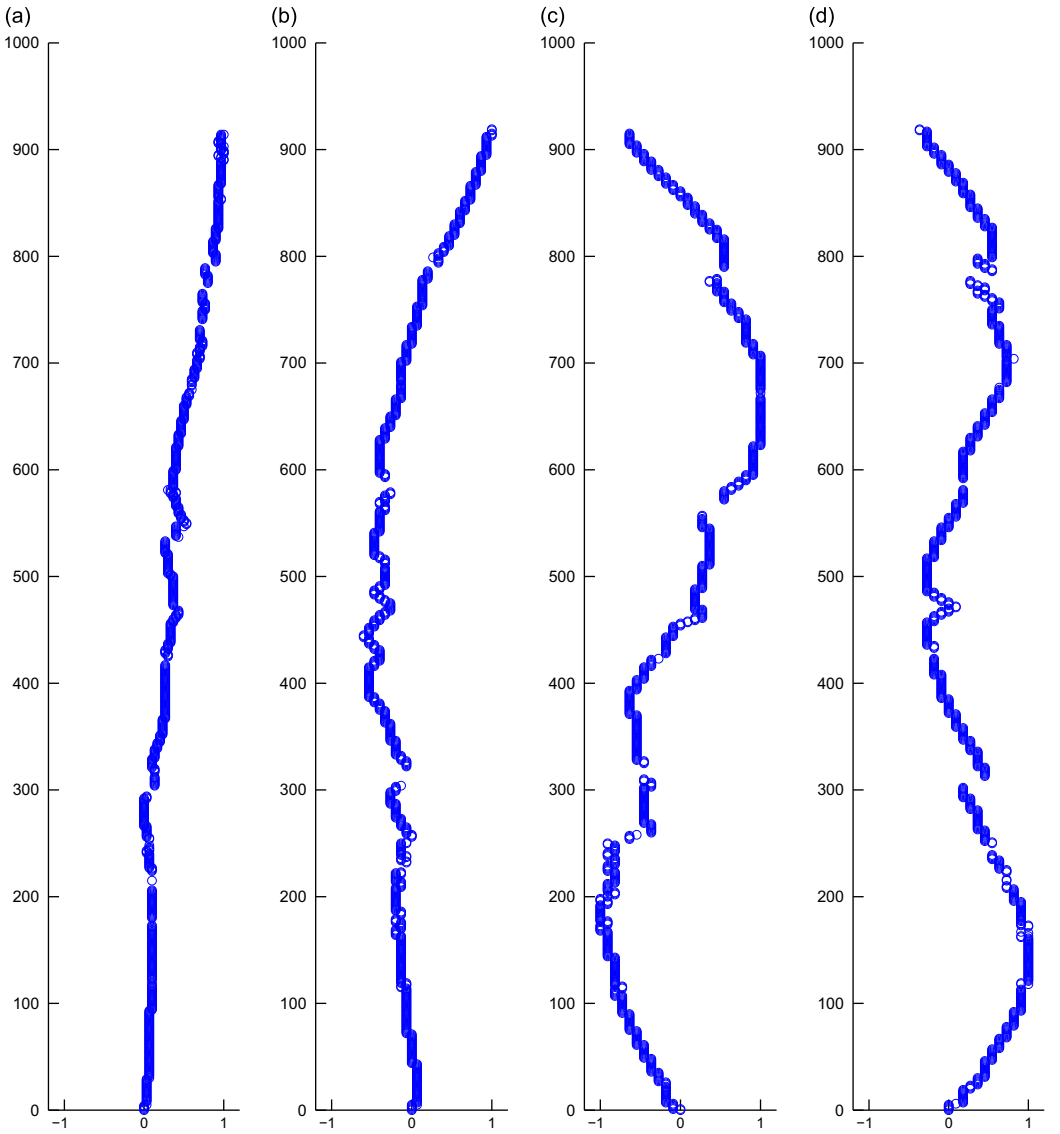

Fig. 9. Operational deflection shapes derived from edge detection of the screenshots from motion magnified video of the cantilever beam: (a) 12.5 Hz, (b) 80 Hz, (c) 226 Hz, and (d) 444 Hz.

5. Pipe test

5.1. Experimental setup

An experiment similar to the instrumented cantilever beam test was performed on a pipe. A section of 4 in schedule 40 PVC pipe was placed on the end of a laboratory bench and held down by hand as shown in Fig. 11(a). The pipe was struck on the free end by an impact hammer as an excitation and the resulting vibrations of the pipe as viewed in the cross-section were recorded with a high-speed camera as shown in Fig. 11(b). The camera recorded video at 20,000 frames per second, with a frame size of 192 by 192 pixels, and in the plane of the pipe the video frame was approximately 147 mm wide and 147 mm tall, with 1.306 pixels per millimeter.

5.2. Visualization of mode shapes with motion magnification

A displacement signal was extracted from the source video cropped to the bottom edge of the pipe, suggesting frequency bands for the motion magnification algorithm. Operational deflection shapes were extracted from the video using the phase-based motion magnification algorithm at 480 Hz, 1200 Hz, and 2400 Hz [14]. The frequency bands used were 475–485 Hz, 1195–1205 Hz, and 2390–2410 Hz with amplification factors of 50, 200, and 400 respectively. Screenshots from the

Table 2
MAC values for the comparison between the accelerometer and edge detected motion magnified video screenshot operational mode shapes.

Accel (Hz)	Edge detected			
	12.5 Hz (%)	80 Hz (%)	226 Hz (%)	444 Hz (%)
12.5	98.26	0.03	18.49	19.20
80	10.65	75.94	0.00	2.75
226	8.15	1.41	84.49	0.10
444	4.45	8.07	0.18	61.27

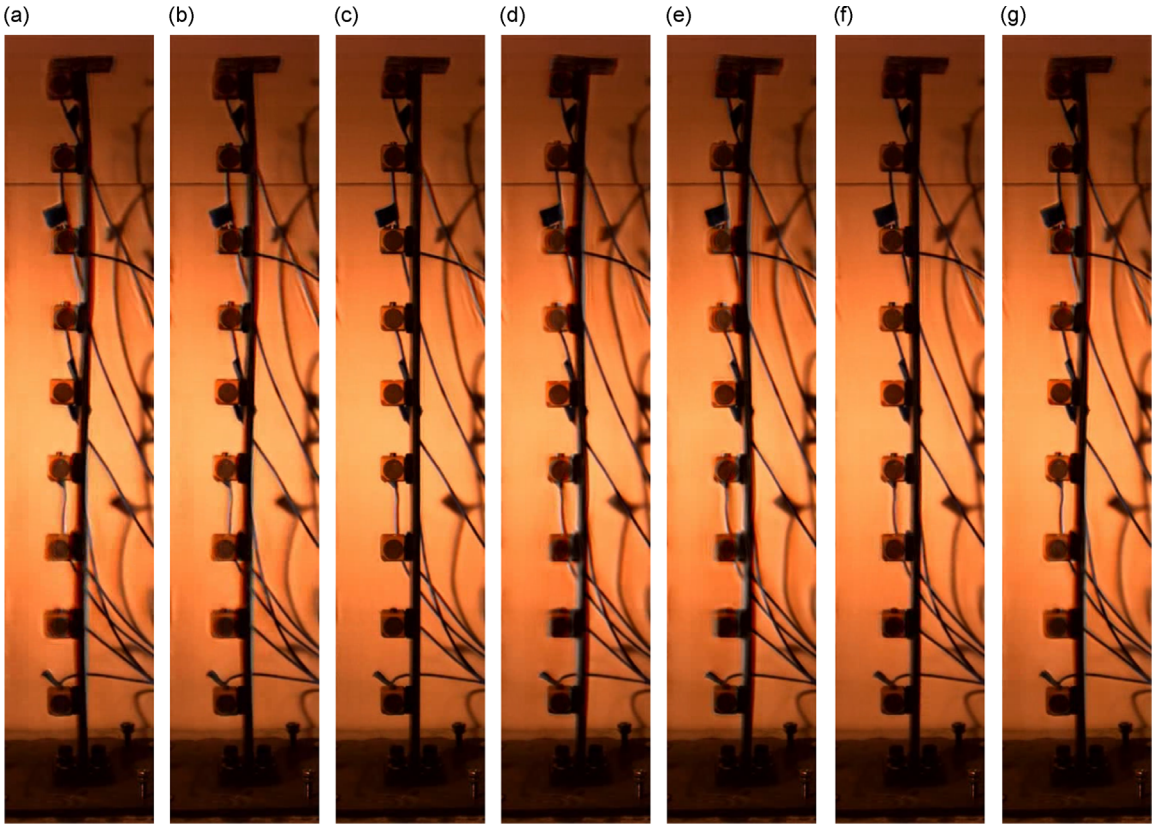

Fig. 10. Series of frames from motion magnified video of the cantilever beam in 222–230 Hz band.

motion magnified video are shown in Fig. 12. These vibrational modes visualized here are the first three circumferential–radial modes of a pipe and match the expected theoretical mode shapes, as described in [26]. The results from the edge detection routine to determine a quantified mode shape from the screenshots are shown in Fig. 13. A series of screenshots from the motion magnified video in the frequency band of 1195–1205 Hz are shown in Fig. 14.

6. Conclusion

In this paper we have demonstrated motion magnification for extracting displacements from high speed video and demonstrated the algorithm's capability of qualitatively identifying the operational deflection shapes of a cantilever beam and a pipe cross section from a video. A verification measurement on a cantilever beam was made and the displacement extracted closely matched those measured by a laser vibrometer. The resulting noise floor of the camera was approximately 1×10^{-5} pixels per root Hertz. The first four operational deflection shapes of the instrumented cantilever beam at 12.5 Hz, 80 Hz, 226 Hz, and 444 Hz were determined from the displacement extracted from cropped regions of the source video and qualitatively visualized using motion magnification. The operational deflection shapes compared favorably to both the expected theoretical mode shapes and the accelerometer data calculated operational deflection shapes with MAC values of 95.58 percent, 97.78 percent, 97.56 percent, and 98.46 percent for the first four modes. Further, the operational deflection shapes were determined directly from the motion magnified video, albeit with lower accuracy. Motion magnification was

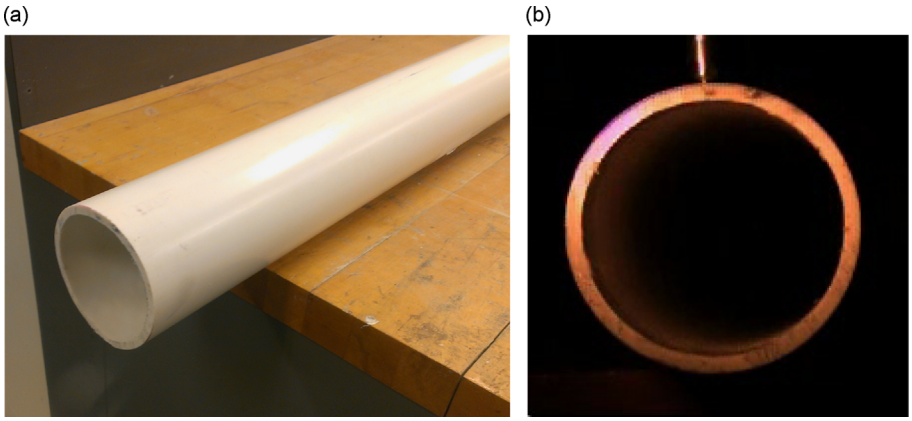

Fig. 11. Pipe measurement: (a) experimental setup and (b) screenshot from video camera.

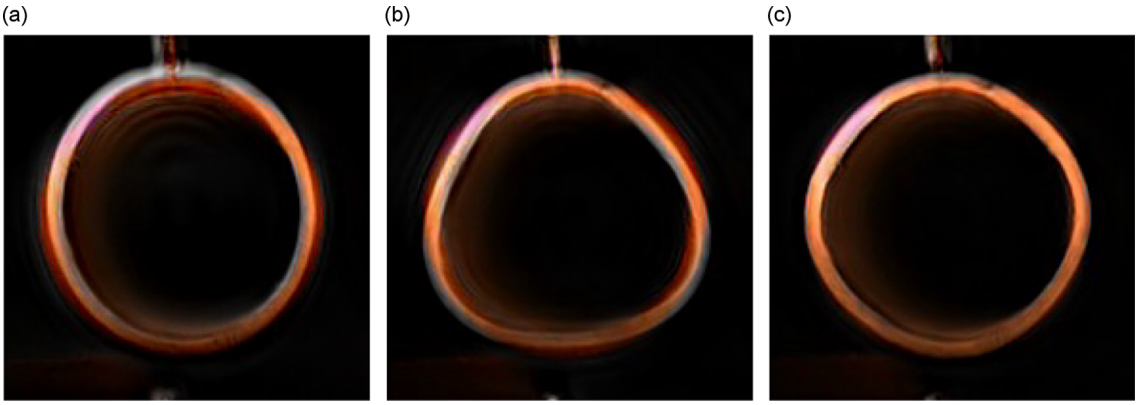

Fig. 12. Screenshots from motion magnified video of pipe cross-section: (a) 480 Hz, (b) 1200 Hz, and (c) 2400 Hz.

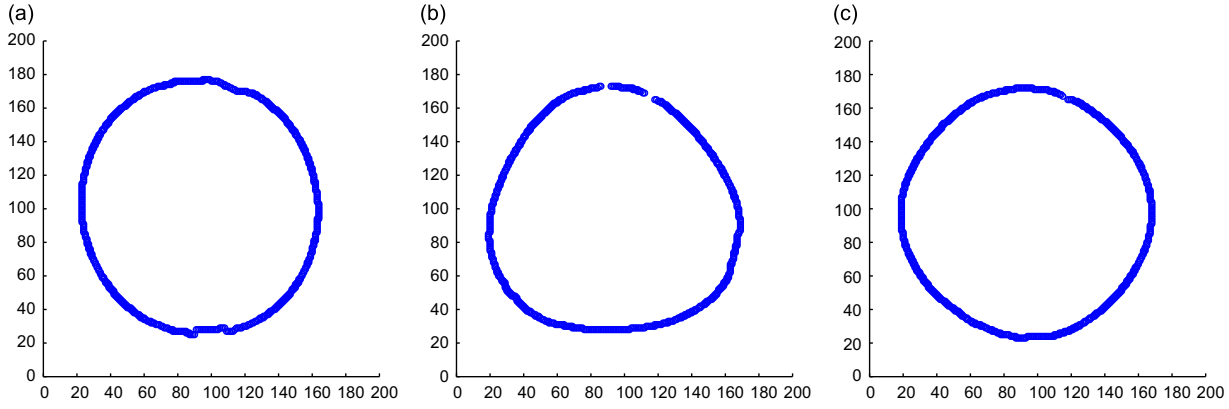

Fig. 13. Operational deflection shapes derived from edge detection of the screenshots from motion magnified video of the pipe cross-section: (a) 480 Hz, (b) 1200 Hz, and (c) 2400 Hz.

also used to visualize and quantify the first three circumferential–radial operational deflection shapes of a 4 in PVC pipe at 480 Hz, 1200 Hz, and 2400 Hz.

Quick visualization of the mode shapes of vibrating structures with a camera is a useful tool for the modal analysis of structures. Measurement of these operational deflection shapes allows for further analysis to be carried out along with quantification of various modal properties that can be used for structural health monitoring purposes. Motion magnification allows for accurate detection and processing of displacements and mode shapes of a vibrating structure for these purposes. The demonstration of the methodology here on two simple structures, a cantilever beam and a pipe, which represent simple models of complex structures that may be found in infrastructure. The methodology will thus be readily adaptable to the

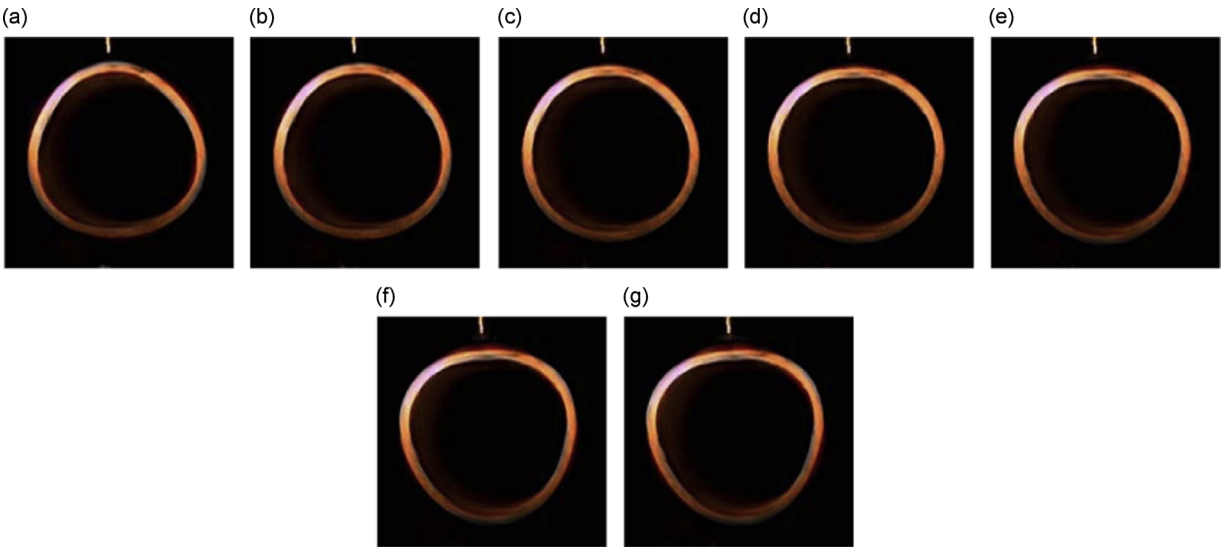

Fig. 14. Series of frames from motion magnified video of the pipe in the 1195–1205 Hz band.

measurement of real-world structures for their condition assessment. The eventual goal is to be able to take video of a structure, measure the displacement, determine its modal characteristics, and analyze those characteristics for any anomalies to detect any damage.

Several areas of this methodology can be improved with further research. Algorithms that automatically extract the operational deflection shapes of coherently vibrating objects in a video would be immensely useful for modal analysis. Extraction of the operational deflection shapes based on the motion magnified video can also be improved by averaging the results from several frames of video and cleaning up the results of the edge detection algorithm. Further challenges need to be overcome for structures with complex geometries, closely coupled modes, and the outdoor measurement of structures, specifically the effects of camera movement, outdoor lighting conditions, and atmospheric aberration due to changes in the index of refraction of air with variations in temperature and wind. Resolving these issues will advance the methodology towards the eventual goal of using the camera as a sensor for the condition assessment of infrastructure.

Acknowledgments

The authors acknowledge the support provided by Royal Dutch Shell through the MIT Energy Initiative, and thank chief scientists Dr. Dirk Smit and Dr. Sergio Kapusta, project managers Dr. Yile Li and Dr. Keng Yap, and Shell-MIT liaison Dr. Jonathan Kane for their oversight of this work. We also acknowledge Draper Laboratory for providing experimental equipment. At the time of this work, Neal Wadhwa was supported by the DoD through the NDSEG fellowship program. Special thanks are due to Reza Mohammadi Ghazi for his help with experimental collection of the data.

Appendix A. Steerable filter taps

Freeman et al. [24] specify steerable filters that are used in this paper. For convenience, their filters are reproduced here in Table A1.

Table A1
Filter coefficients to compute horizontal and vertical local phase and local amplitude.

Tap #	G_{f1}	G_{f2}	H_{f1}	H_{f2}
−4	0.0094	0.0008	−0.0098	0.0008
−3	0.1148	0.0176	−0.0618	0.0176
−2	0.3964	0.1660	0.0998	0.1660
−1	−0.0601	0.6383	0.7551	0.6383
0	−0.9213	1.0000	0.0000	1.0000
1	−0.0601	0.6383	−0.7551	0.6383
2	0.3964	0.1660	−0.0998	0.1660
3	0.1148	0.0480	0.0618	0.0176
4	0.0094	0.0008	0.0098	0.0008

Filter	Filter in x	Filter in y
Real horizontal (G_2^0)	G_{f1}	G_{f2}
Imaginary horizontal (H_2^0)	H_{f1}	H_{f2}
Real vertical ($G_2^{\pi/2}$)	G_{f2}	G_{f1}
Imaginary vertical ($H_2^{\pi/2}$)	H_{f2}	H_{f1}

References

[1] M. Ashory, Correction of mass-loading effects of transducers and suspension effects in modal testing, *Proceedings of the 16th International Modal Analysis Conference*, vol. 3243, 1998, p. 815.
[2] O. Cakar, K. Sanliturk, Elimination of transducer mass loading effects from frequency response functions, *Mechanical Systems and Signal Processing* 19 (1) (2005) 87–104.
[3] C.R. Farrar, T.W. Darling, A. Migliori, W.E. Baker, Microwave interferometers for non-contact vibration measurements on large structures, *Mechanical Systems and Signal Processing* 13 (2) (1999) 241–253.
[4] A. Stanbridge, D. Ewins, A. Khan, Modal testing using impact excitation and a scanning LDV, *Shock and Vibration* 7 (2) (2000) 91–100.
[5] H. Van der Auweraer, H. Steinbichler, S. Vanlanduit, C. Haberstok, R. Freymann, D. Storer, V. Linet, Application of stroboscopic and pulsed-laser electronic speckle pattern interferometry (ESPI) to modal analysis problems, *Measurement Science and Technology* 13 (4) (2002) 451.
[6] S. Patsias, W. Staszewskiy, Damage detection using optical measurements and wavelets, *Structural Health Monitoring* 1 (1) (2002) 5–22.
[7] A.M. Wahbeh, J.P. Caffrey, S.F. Masri, A vision-based approach for the direct measurement of displacements in vibrating systems, *Smart Materials and Structures* 12 (5) (2003) 785.
[8] J.J. Lee, M. Shinozuka, A vision-based system for remote sensing of bridge displacement, *NDT&E International* 39 (5) (2006) 425–431.
[9] E. Caetano, S. Silva, J. Bateira, A vision system for vibration monitoring of civil engineering structures, *Experimental Techniques* 35 (4) (2011) 74–82.
[10] M.N. Helfrick, C. Niezrecki, P. Avitabile, T. Schmidt, 3D digital image correlation methods for full-field vibration measurement, *Mechanical Systems and Signal Processing* 25 (3) (2011) 917–927.
[11] A. Kolb, E. Barth, R. Koch, R. Larsen, Time-of-flight sensors in computer graphics, *Proceedings of Eurographics* (*State-of-the-Art Report*), 2009, pp. 119–134.
[12] C. Liu, A. Torralba, W.T. Freeman, F. Durand, E.H. Adelson, Motion magnification, *ACM Transactions on Graphics* 24 (2005) 519–526. URL ⟨http://doi.acm.org/10.1145/1073204.1073223⟩.
[13] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, W. Freeman, Eulerian video magnification for revealing subtle changes in the world, *ACM Transactions on Graphics* (*Proceedings of SIGGRAPH*), Vol. 31.
[14] N. Wadhwa, M. Rubinstein, F. Durand, W.T. Freeman, Phase-based video motion processing, *ACM Transactions on Graphics* (*Proceedings of SIGGRAPH 2013*), Vol. 32, no. 4.
[15] M. Rubinstein, Analysis and Visualization of Temporal Variations in Video, PhD Thesis, Massachusetts Institute of Technology, February 2014.
[16] N. Wadhwa, M. Rubinstein, F. Durand, W.T. Freeman, Riesz pyramid for fast phase-based video magnification, *2014 IEEE International Conference on Computational Photography* (*ICCP*), Santa Clara, CA, May 2-4, 2014, IEEE, 2014.
[17] J. Chen, N. Wadhwa, Y.-J. Cha, F. Durand, W. Freeman, O. Buyukozturk, Structural modal identification through high speed camera video: motion magnification, in: J. De Clerck (Ed.), *Topics in Modal Analysis I*, Vol. 7, *Conference Proceedings of the Society for Experimental Mechanics Series*, Springer International Publishing, 2014, pp. 191–197.
[18] E.P. Simoncelli, W.T. Freeman, The steerable pyramid: a flexible architecture for multi-scale derivative computation, *Proceedings of the 1995 International Conference on Image Processing (ICIP '95)*, Vol. 3, IEEE Computer Society, Washington, DC, USA, 1995, pp. 444–447. URL ⟨http://dl.acm.org/citation.cfm?id=839284.841271⟩.
[19] L. Hermans, H. van der Auweraer, Modal testing and analysis of structures under operational conditions: industrial applications, *Mechanical Systems and Signal Processing* 13 (2) (1999) 193–216.
[20] B. Horn, B. Schunck, Determining optical flow, *Artificial Intelligence* 17 (1–3) (1981) 185–203.
[21] B.D. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, *Proceedings of the Seventh International Joint Conference on Artificial Intelligence (IJCAI '81)*, 1981, pp. 674–679.
[22] D.J. Fleet, A.D. Jepson, Computation of component image velocity from local phase information, *International Journal of Computer Vision* 5 (1) (1990) 77–104, http://dx.doi.org/10.1007/BF00056772.
[23] T. Gautama, M. Van Hulle, A phase-based approach to the estimation of the optical flow field using spatial filtering, *IEEE Transactions on Neural Networks* 13 (5) (2002) 1127–1136.
[24] W.T. Freeman, E.H. Adelson, The design and use of steerable filters, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13 (9) (1991) 891–906.
[25] K. Worden, Data processing and experiment design for the restoring force surface method, Part I: Integration and differentiation of measured time data, *Mechanical Systems and Signal Processing* 4 (4) (1990) 295–319.
[26] R.D. Blevins, *Formulas for Natural Frequency and Mode Shape*.