

FULL-FIELD VIBRATION MEASUREMENT USING CAMERA THROUGH SHEARLET SYSTEM

C. A. Wee^{1*}, M. Z. Ngali¹

¹Faculty of Mechanical and Manufacturing Engineering,
Universiti Tun Hussein Onn Malaysia,
*Email: gd160056@siswa.uthm.edu.my

ABSTRACT

Traditionally, vibration measurement is done using an accelerometer or Laser Doppler Vibrometer, which is an intrusive and single point measurement respectively. This paper demonstrated vibration signal extracted from a perceptually invisible vibrating object using only a camera non-intrusively at full-field capacity. The camera that is capable to capture 1000 frames-per-second video was used to capture the motion of the vibrating object. Each frame of the video was decomposed using complex shearlet transform and the extracted signal was compared to an accelerometer. Shearlet decomposed each frame of the video into complex coefficients which were later used to recover the motion between two consecutive frames. Phase information that were used to retrieve the vibration signal were weighted to decrease unwanted noise. Resonant frequencies of a simply supported beam at 104.0 Hz, 209.5 Hz and 396.0Hz were successfully recovered. Single frequency extraction from an excited loud-speaker at specific frequencies were also conducted and showed a clear peak-to-valley frequency spectrum recovery. The potential of using camera as a full-field displacement measurement where each pixel acting as a vibrometer was explored. Full-field test to recover the mode shape of a circular membrane showed promising result with eight mode shapes successfully retrieved. The experiments proved that each individual pixel was able to retrieve motion at subpixel level that is at 0.00001 pixel scale.

Keywords: full-field, vibration measurement, camera, mode shape, shearlet, subpixel motion.

INTRODUCTION

Overview

Vibration of an object is traditionally measured using accelerometer that requires the sensor attached to the vibrating body. While accelerometer is accurate and precise in measuring displacement, it is intrusive to the measuring object or structure itself as the mass of the accelerometer somehow altered the dynamic response. There are devices that uses non-intrusive method to overcome this problem. Laser Doppler Vibrometer (LDV) projected a laser beam onto the surface of a vibrating surface and measure its velocity response from Doppler shift effect[1]. It has high spatial resolution and could be easily set-up. However, LDV is a single point axial vibrometer as it only measure one point at a time. Continuous scanning laser Doppler vibrometer (CSLDV) have been introduced to overcome the limitation of LDV. CSLDV employed the LDV whose laser beam is sweeping across the surface of vibrating object to measure the vibration at multiple points.

Recently, advanced full-field measurement methods had been introduced [2-4] by using high speed camera. These techniques involved using a high frame rate camera to capture a series of images at thousandth of second per frame. Each frame is then decomposed into Fourier domain using complex steerable pyramid (CSP) [5] at multi-resolution and multi-orientations. The phase of the decomposed image was then processed and filtered to reduced noise and extracted the encoded vibration signal in its phase. This method is highly efficient and accurate in extracting sub-pixel motion as compared to optical-flow method such as Lukas-Kanade[6] and Horn-Schunck[7] which tends to be affected by the noise as estimation operations were performed spatially. In contrast, CSP operations are all in Fourier domain.

Despite quite a number of studies had been done on the full-field measurement method, it still has many potential to be improved and explored. This paper is thus written to address the lacking. A more sophisticated method in decomposing the image using Shearlet transformation is introduced to improve the vibration signal extraction from a subpixel motion video.

This text is organized as follows: Section 2 we discussed the basic of phase-based motion estimation. Section 3 explained wavelet and shearlet system. Section 4 discussed the experiments set-up. Section 5 showed experiment result in validating the proposed method and obtaining the modal shape of a circular membrane.

Phase-Based Motion Estimation and Shearlet System

Phase-based motion estimation is a method that employ image processing technique to decompose an image into phase and magnitude domain using wavelet basis functions. Wavelet bases are well localized in space as opposed to Fourier bases that sparse in representation[8] and thus it is suitable to be used in full-field displacement measurement. It is not to be confused with phase correction [9] which estimates the motion of two images globally by computing the Phase-Only-Correlation (POC) and finding the peak of shifting using cross-power spectrum.

Well established wavelet image decomposition is steerable pyramid[5]. In [10], the authors used CSP to magnify motion in a video that is invisible to naked human eyes, while *Davis et al.* [3] successfully extended the work to recover sounds from high speed video. Steerable pyramid uses steerable Gabor-like wavelets as the basis function to represent image coefficients in multi-resolution and multi-orientations. In complex steerable pyramid, the analytic signal is made up from complex-valued sinusoids [11]. Real and imaginary parts of the coefficients have more advantage over real-valued steerable pyramid. For instances, the magnitude of a complex coefficients are shift-invariance and its phase contain more image features especially edge. The phase in complex coefficients near an edge varies linearly with its distance to the edge[12].

Mathematically, CSP decomposed image I into different complex-valued spatial sub-bands $S_{r,\theta}$ corresponding to scale r and orientation θ . The basis functions during this transformation are scaled and oriented accordingly. At each sub-band, a complex-valued image is defined in terms of amplitude A and phase φ , by abusing of notation

$$A(r, \theta, x, y)e^{i\varphi(r, \theta, x, y)} \quad (1)$$

To extract the vibration signal between two consecutive images, local phase different between them is calculated. The local phase different φ_d between an image at time t and a reference image t_0 is defined as

$$\begin{aligned} \varphi_d(r, \theta, x, y, t) &= \varphi(r, \theta, x, y, t) \\ &\quad - \varphi(r, \theta, x, y, t_0) \end{aligned} \quad (2)$$

However such operations will involve the phase different of unwanted noise too. Fortunately, the amplitude from CSP provide us the details about the noise. High amplitude corresponds to low noise and vice-versa. Making use of such information, we were able to filter and weight the noise. It is worth to mentions that such operation is valid for small motion [13], which this paper is focused on. Frames were then decomposed using shearlet transform. It was proved that it achieved better result than the conventional wavelet transform [14]. Wavelet tends to treat different axes equally, such property resulted in poor anisotropic feature detection [15]. Since image is a 2D signals, such behavior will have impact on the image decomposition of highly texturized image with a lot of edges. This is where shearlet system outperform the wavelet as it is highly scalable and shear-able. However classic shearlet suffer from ineffectiveness of computation when the shear parameter is large. In modern cone-adapted shearlet system [15], frequency domain is separated into two region: low-pass residual at center and two conic regions to overcome the problem of large shearing parameters as shown on the right of Figure 1.

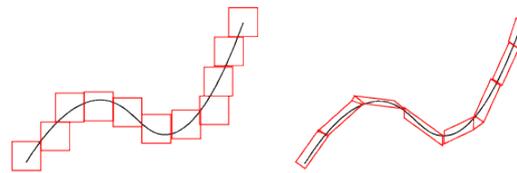


Figure 1. (left) Isotropic and; (right) Anisotropic feature approximation.

Large shearing occur when the line is near to horizontal, without cone-adapted method, multiple times of transformation are needed to fit the line perfectly. With the modern shearlet system, complex coefficients of an image in level r , cone k and orientation θ is therefore defined as

$$A(r, k, \theta, x, y, t) e^{i\varphi(r, k, \theta, x, y, t)} \quad (3)$$

and its respective phase different is

$$\varphi_a(r, k, \theta, x, y, t) = \varphi(r, k, \theta, x, y, t) - \varphi(r, k, \theta, x, y, t_0) \quad (4)$$

Figure 2 showed that shearlet system is well localized compared to steerable pyramid. Image discontinuities such as the edge of the umbrella could be seen clearly in the shearlet decomposed phase image.

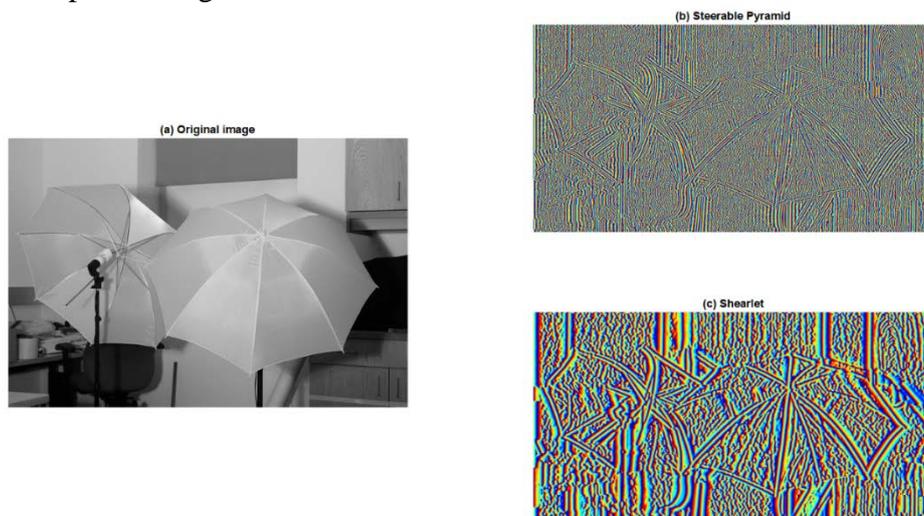


Figure 2. (a) Umbrella image [16]. Phase image decomposed using (a) CSP and (b) shearlet.

EXPERIMENTS

Employing equation (4) to calculate motion signal, we were able to extract the vibration signal for a freely hanged steel bar as shown in Figure 3. However, equation (4) is not noise resistant. Fortunately, the amplitude $A(r, k, \theta, x, y, t)$ from shearlet coefficients provided us with the details of noise. Low image contrast usually have high noise of phase. Amplitude of the image coefficients is a simple measurement of image contrast, where high amplitude indicates low image contrast and vice-versa. Thus, high amplitude corresponds to low noise. With such information, we were able to filter and weight the noise in the phase.

The camera was pointed perpendicular to the steel bar to capture all vertical motion produced by the steel bar. Two seconds of video at 1000 frame-per-second at an effective pixel of 1136x384 were recorded. A total of 2000 frames were generated from the video and stored in Bitmap (BMP) format to conserve its original image details. Approximately 12 gigabytes of data were generated from a single video alone. The frequency limit that we were able to extract from this video is 500Hz, in accordance to the Nyquist Theorem.

In order to obtain a relative motion between two frames, one could use a reference frame from a still image or by averaging the image intensities of a video over time. In our paper, we have chosen the later. Since it is not always possible to obtain a still reference image of a vibrating object in real life, thus it would be more practical to obtain a reference frame of averaged intensities over time. Besides, by averaging the frames over time, the final image would be less disturbed by noise. Such method is possible as the motion considered in this paper is at subpixel level and averaging the intensities will not change the average intensity of a pixel.

As shearlet system decomposed the frame into multi-resolution and multi-orientation, we decided to use the finest level of the coefficients as it contains more high frequency information. As for the orientation, we have chosen the direction that is parallel to the motion of the vibrating object.

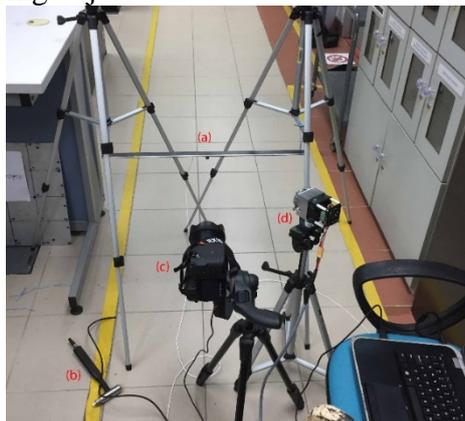


Figure 3. (a) Freely hanged steel bar on both sides using thin ropes. (b) Impact hammer used to subtly excite the steel bar. (c) Sony RX10M3 camera. (d) LED to light up the recording scene.

Next, we tested our method in a signal frequency scenario. A loud speaker is excited to produce a frequency at 225Hz, 300Hz, 425Hz and 475Hz. The amplitude is set to as lowest as possible so as to produce a motion that is subtle and invisible to human eyes. The camera was pointed fronto-parallel to the speaker so that the motion of the diaphragm is moving out of the image to the right.

Finally, we proceed with the full-field testing of camera as a multi-sensor vibrometer. A piece of latex membrane with stripes was mounted onto a PVC tube of diameter 15cm, as shown in Figure 4. The membrane is then placed in front of a loud speaker to be excited.



Figure 4. (left) A frame from the captured video and; (right) piece of latex membrane mounted on a PVC tube.

RESULTS AND DISCUSSION

The peaks of frequency spectrum obtained from our method matched closely with the value from accelerometer, as shown in Figure 5. Three resonant frequencies of the steel bar had been recovered successfully at 104.0Hz, 209.5Hz and 396.0Hz. However, resonant frequency at 347.7Hz and 356.0Hz were buried in noise in our measurement.

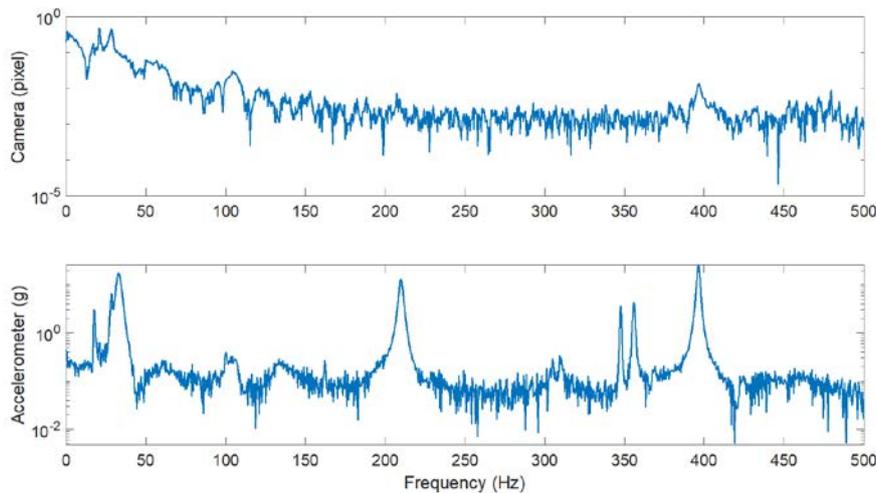
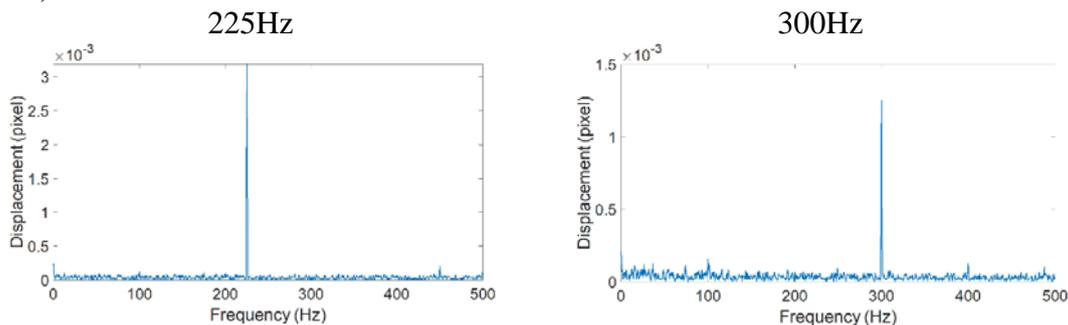


Figure 5. Frequency spectrum for camera-based and accelerometer measurement.

For the peaks exist below 50Hz, we verified that it was the unavoidable vibration that were caused by uncontrollable surrounding, such as the ventilation system.

In a single frequency test, our method was able to extract the frequency in clear peak as shown in Figure 6. This showed that our method have high signal-to-noise ratio (SNR).



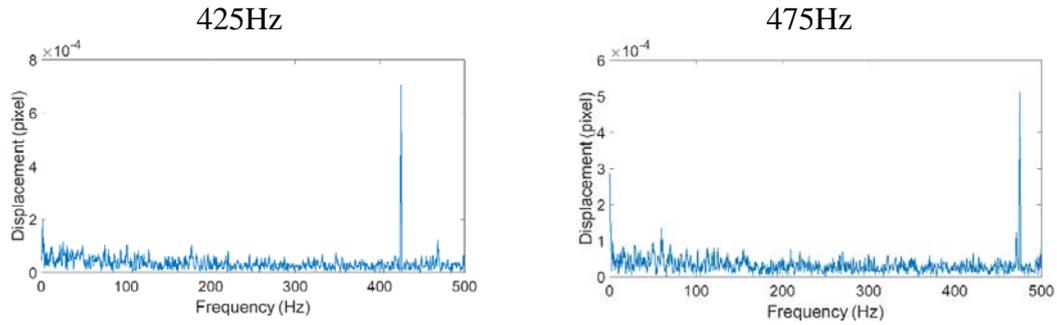


Figure 6. Frequency spectrum for single frequency test.

For our full-field test, we were able to recover eight different mode shapes at different frequencies as shown in Figure 7. The displacement of the membrane is lower than 0.00003 pixel. The conversion factor for pixel displacement to millimeter we used was 0.12 mm/pixel.

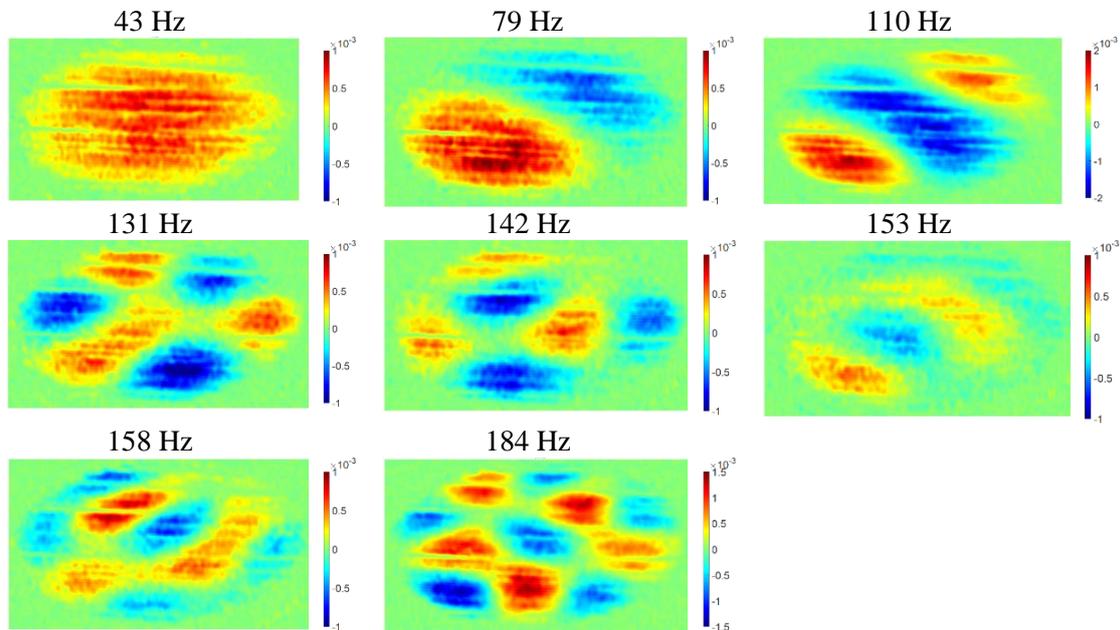


Figure 7. Recovered mode shapes at pixel displacement scale.

CONCLUSION

This research demonstrated that camera could be used as a full-field vibrometer to measure tiny vibration that is perceptually invisible to human eyes. Every pixel in the camera sensor acted as an individual vibrometer sensor. By using shearlet image decomposition, our algorithm was able to extract sub-pixel level motion that is at 0.00001 pixel scale. We successfully showed that the mode shape of a membrane at multiple frequencies were retrieved using the method we proposed and proved our statement that a camera could be used as full-field vibrometer.

ACKNOWLEDGEMENT

This research is supported by the Fundamental Research Grant Scheme (FRGS), Grant No: FRGS/2/2015 Vot 1543 by Ministry of Higher Education, Malaysia.

REFERENCES

- [1] P. Castellini, M. Martarelli, and E. Tomasini, "Laser Doppler Vibrometry: Development of advanced solutions answering to technology's needs," *Mechanical Systems and Signal Processing*, vol. 20, no. 6, pp. 1265-1285, 2006.
- [2] A. Davis, K. L. Bouman, J. G. Chen *et al.*, "Visual vibrometry: Estimating material properties from small motions in video." pp. 5335-5343.
- [3] A. Davis, M. Rubinstein, N. Wadhwa *et al.*, "The visual microphone: passive recovery of sound from video," 2014.
- [4] D. A. Ehrhardt, M. S. Allen, S. Yang *et al.*, "Full-field linear and nonlinear measurements using continuous-scan laser doppler vibrometry and high speed three-dimensional digital image correlation," *Mechanical Systems and Signal Processing*, vol. 86, pp. 82-97, 2017.
- [5] E. P. Simoncelli, and W. T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation." pp. 444-447.
- [6] B. D. Lucas, and T. Kanade, "An iterative image registration technique with an application to stereo vision," 1981.
- [7] B. K. Horn, and B. G. Schunck, "Determining optical flow," *Artificial intelligence*, vol. 17, no. 1-3, pp. 185-203, 1981.
- [8] S. Mallat, *A wavelet tour of signal processing*: Academic press, 1999.
- [9] Y. Yu, and J. Wang, "Highly Accurate Estimation of Sub-pixel Motion Using Phase Correlation." pp. 186-193.
- [10] N. Wadhwa, M. Rubinstein, F. Durand *et al.*, "Phase-based video motion processing," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 4, pp. 80, 2013.
- [11] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE signal processing magazine*, vol. 22, no. 6, pp. 123-151, 2005.
- [12] R. Anderson, N. Kingsbury, and J. Fauqueur, "Determining multiscale image feature angles from complex wavelet phases." pp. 490-498.
- [13] T. Gautama, and M. Van Hulle, "A phase-based approach to the estimation of the optical flow field using spatial filtering," *IEEE Transactions on Neural Networks*, vol. 13, no. 5, pp. 1127-1136, 2002.
- [14] H. K. Sevindir, and C. Yazıcı, "Comparison of Wavelet and Shearlet Transforms for Medical Images," *Appl. Math*, vol. 10, no. 4, pp. 1447-1452, 2016.
- [15] G. Kutyniok, and D. Labate, "Introduction to shearlets," *Shearlets*, pp. 1-38: Springer, 2012.
- [16] D. Scharstein, H. Hirschmüller, Y. Kitajima *et al.*, "High-resolution stereo datasets with subpixel-accurate ground truth." pp. 31-42.