



Smaller Than the Eye Can See: Vibration Analysis with Video Cameras

Oral BUYUKOZTURK¹, Justin G. CHEN¹, Neal WADHWA¹, Abe DAVIS¹,
Frédo DURAND¹, William T. FREEMAN¹

¹ Massachusetts Institute of Technology
77 Massachusetts Ave.

Cambridge, MA 02139, USA

Contact e-mail: obuyuk@mit.edu

Abstract. In the field of non-destructive testing the tools that we use to conduct inspections has evolved over the years. Originally, inspectors used their senses: sight, hearing, and touch to evaluate structures, but now they use equipment such as boroscopes, dye penetrant, microphones, and ultrasonic transducers, which can augment the human senses and/or make much finer quantitative measurements beyond any human capabilities. Recent developments in computer vision have created a “motion microscope”, where small imperceptible motions in videos can be magnified and easily seen; this technique is called motion magnification. Additionally, the method enables quantitative analysis of structures from video as the information acquired for motion magnification also provides displacement signals as a basis for vibrational analysis.

In this paper, we present applications of motion magnification to vibration analysis in non-destructive testing. We describe the methodology and theory behind motion magnification, the workflow from video to modal information, and discuss strategies for processing the numerous signals collected from video data. We demonstrate making video measurements of small structures and rotating machinery in controlled conditions, and in near real time, extracting displacements from video and doing simple modal and frequency analysis. We also present the results of outdoor measurements of civil infrastructure and discuss future applications for the methodology in non-destructive testing.

Introduction

The origins of non-destructive testing (NDT) are as simple as visual and acoustic inspection of objects by one’s own eyes and ears. As technology has improved, inspectors have made use of a wide range of measurement tools to augment their senses and quantify the presents or effects of damage on a structure. Developments in computer vision have made it possible to qualitatively and quantitatively analyze small, imperceptible motion in camera videos. This capability allows for passive non-contact measurement of vibrations and displacements in structures, potentially at low cost as cameras are less expensive than laser-based, radar, or radiographic techniques. This can help with modal analysis and vibration based NDT techniques or qualitative inspection looking for parts of a structure that vibrate excessively.



In this paper, we present applications of motion magnification to vibration analysis in NDT and structural health monitoring (SHM). Section 1 describes the methodology behind motion magnification, specifically how videos are processed to measure the vibrations of objects in the video. Several applications of video measurements are presented in Section 2, for small structures, rotating machinery, and civil infrastructure. Section 3 discusses uses and limitations of the methodology, and potential future improvements for NDT applications.

1. Video Processing for Vibration Measurement

This section is a brief summary of how video processing is accomplished for vibration measurement, both for motion magnification to visualize imperceptible motions in videos, and for quantitative measurement of displacements and vibrations. The full descriptions for the processing procedures can be found in the following cited papers for readers requiring more detail [1-6].

1.1 Motion Magnification

Motion Magnification is an algorithm for the amplification of small motions in videos, acting as a sort of “motion microscope”. The basic principle is to obtain a representation for the video such that signals representing the motions of objects in the video which can be processed 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 [1] decomposes the signal of a video into the local spatial amplitude and phase using a complex-valued steerable pyramid filter bank [7]. The local spatial phase signals, representing motion, are Fourier decomposed into a series of sinusoids representing harmonic motion in the time domain. The phase signals are then temporally bandpass filtered, amplified, and recombined to form a motion magnified video. The result is a video with motion in the specified band of temporal frequencies amplified and made more visible. This procedure can be run in real-time for reasonably sized videos and framerates with a slightly different processing procedure [2].

1.2 Displacement Measurement

This procedure can be “short-circuited” and the phase information from the video decomposition can be used to provide displacement information of structures. The spatial local phase can be used to calculate the displacement of objects [8, 9]. Only regions with sufficient local contrast can provide accurate motion information, as textureless regions do not provide any information on the motion of the object. Pixels with sufficient local contrast are identified and a set of representative motion signals is determined for the video. This set of displacement time series can be used for modal analysis or vibration analysis of an object in the video.

2. Example Applications

In this section we discuss several example applications of our methodology including a lightweight structure, rotating machinery, civil infrastructure, and a material test.

2.1 Lightweight Beams

Lightweight structures are a particularly relevant application for non-contact measurement

methods because the attachment of traditional sensors such as accelerometers or strain gauges may alter the system dynamics by adding mass or stiffness. As a part of a separate experimental study, we made measurements of a series of lightweight cantilever beams made of different plastic materials to look for differences in the material properties. This was quantified by looking at the frequency of the first resonant modes. An example measurement is presented here with the desktop experimental setup shown in Fig. 1a, and a screenshot of the recorded video shown in Fig. 1b.

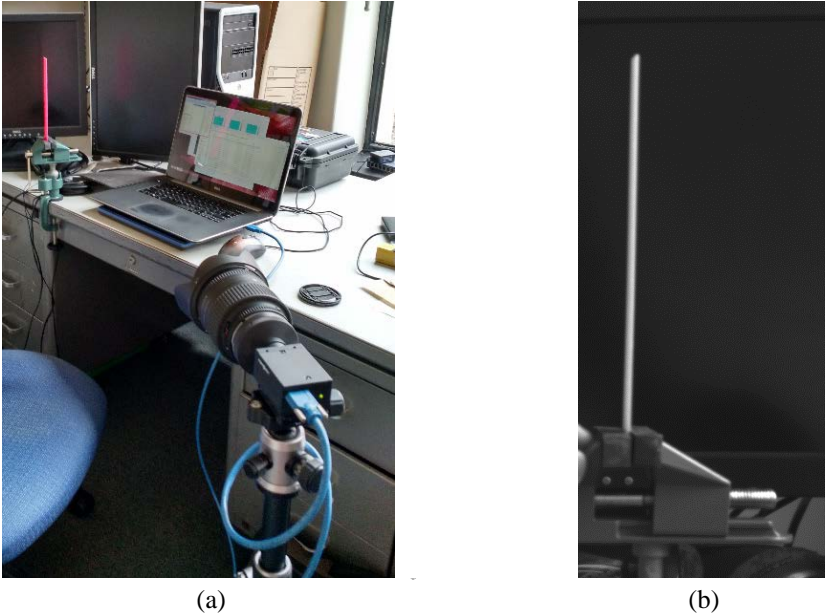


Fig. 1. (a) Experimental setup and (b) screenshot from the recorded video

The input video had a resolution of 1200 x 512 pixels, 1000 frames recorded at 161.4 frames per second (fps). From this video, the first two resonant modes of the beam were measured at 11.95 Hz and 78.94 Hz, as shown in Fig. 2a, and the frequency spectrum is shown in orange in Fig. 2b. Measurements with an accelerometer were also done, with the frequency spectra shown in blue and red in Fig. 2b for two different locations of the accelerometer on the beam. They show a significant shift in the frequency of the first resonant mode, due to the fact that the accelerometer mass of 16 grams is of similar mass as the beam itself at 10 grams.

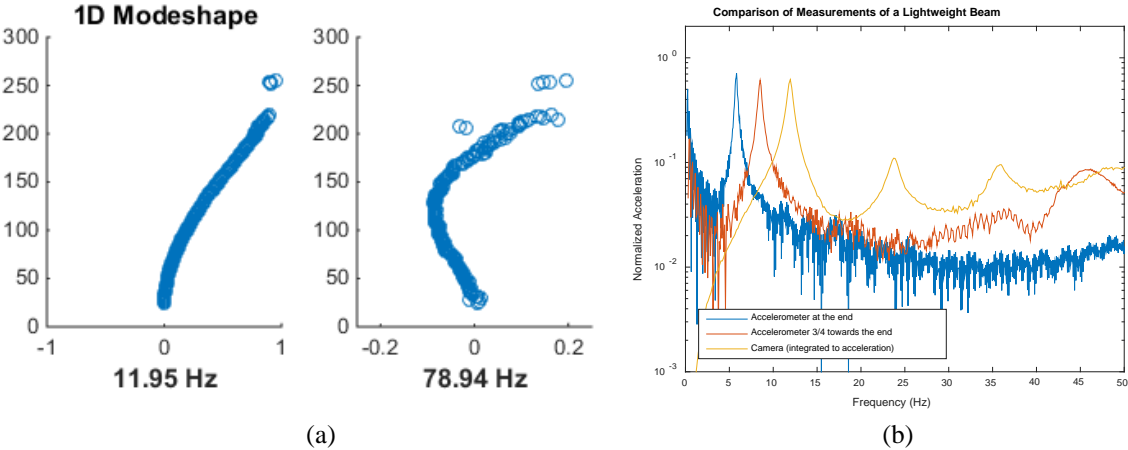


Fig. 2. Measured (a) mode shapes and (b) frequency spectra

2.2 MIT Facilities Measurement

A camera measurement was made of a motor that runs a water pump in the physical facilities of the Massachusetts Institute of Technology (MIT). A picture of the experimental setup is shown in Fig. 3a with a screenshot of the recorded video, 1280 x 720 pixels at 2000 fps, shown in Fig. 3b. Two regions of interest were chosen for extraction of displacements in the video and calculation of the measured frequency spectra. Additionally, a laser vibrometer was used to separately measure the motions of the motor and the ground as a reference for the vibrations in the environment itself, caused by the motor and other machinery in the area.



Fig. 3. Camera measurement of a motor running a pump, (a) experimental setup and (b) screenshot from video with regions of interest 1 and 2

The frequency spectra extracted from the regions of interest 1 and 2 in Fig. 3b are shown in blue in Fig. 4a and 4b respectively. The frequency spectra reasonably match for lower frequencies, with those measured for some of the resonant frequencies measured by the laser vibrometer, shown in red. There are some issues for a camera measurement in an environment with many sources of vibration as evidenced by the laser vibrometer measurement of the ground. There are many vibrations in the environment that shake the camera, giving the appearance of motion of the object. Reference background objects included in the video frame can be used to measure and rule out motion due to global motion of the camera.

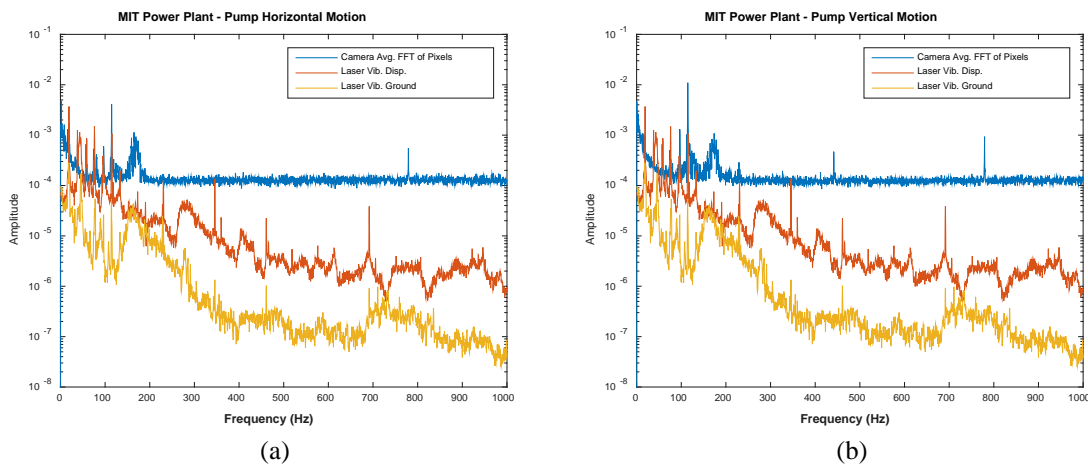


Fig. 4. Camera measured frequency spectra with comparison to laser vibrometer measurement for (a) horizontal motion at location 1, and (b) vertical motion at location 2

2.3 Taipei 101 Earthquake

In this application, we demonstrate the approach on a video from YouTube which is not recorded by the researchers, for measurement of civil infrastructure. The example video is from a fixed camera within the observation deck of the Taipei 101 tower in Taipei, Taiwan during a M6.4 earthquake which occurred at a distance of 120.85 km south east, on 2015/04/20 at 01:42:58 UTC [10]. The source normal framerate (30 fps) video starts at 01:43:01 UTC, 3 seconds after the earthquake starts, and continues for 2 minutes. The Taipei 101 tower is unique in that the structure's tuned mass damper, a massive steel pendulum, is plainly visible in the indoor observation desk, and also this video. A screenshot of the video is shown in Fig. 5a. The displacement of the tuned mass damper is extracted from this video sequence, from the cropped video of size 640 x 130 pixels, shown in Fig. 5b. The mask for pixels with satisfactory displacement information is shown in Fig. 5c.



Fig. 5. (a) Screenshot from source video, (b) video cropped to tuned mass damped, and (c) pixels with valid displacements

The displacement extracted from the tuned mass damper in the video is shown in Fig. 6. At 43 seconds after the video starts, there is clearly a large amount of motion, also visible in the video, that likely corresponds to the arrival of S-waves from the earthquake. At 28 seconds, there is a much smaller disturbance not visible in the video, but revealed by the processing which corresponds to the arrival of P-waves from the earthquake. This behavior is standard for earthquakes as the wave speed of P-waves is faster than those of S-waves.

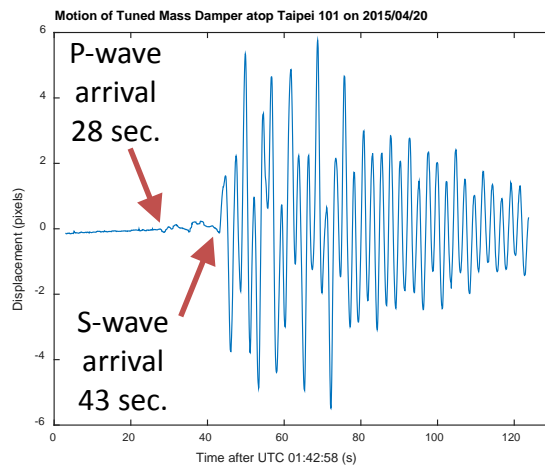


Fig. 6. Extracted motion of tuned mass damper from video

2.4 MIT Green Building Antenna

This application is a measurement of the Green Building at the Massachusetts Institute of Technology over a long distance, originally presented at the NDT-CE 2015 conference [6]. The video measurement of the Green Building was made from a distance of approximately 175 m (by land) on a terrace of the Stata Center, another building at MIT as shown in Fig. 7a. A picture of the experimental setup and the view from the measurement location is shown in Fig. 7b. A camera was used to record a 150 second long video at 10 fps and a resolution of 1200×1920 , where 1 pixel corresponds to about 3.65 cm at the depth of the structure. A screenshot from the recorded video is shown in Fig. 7c showing the Green Building filling most of the frame of the video, with the crow's nest or antenna tower visible as the long thin structure on the roof of the building.

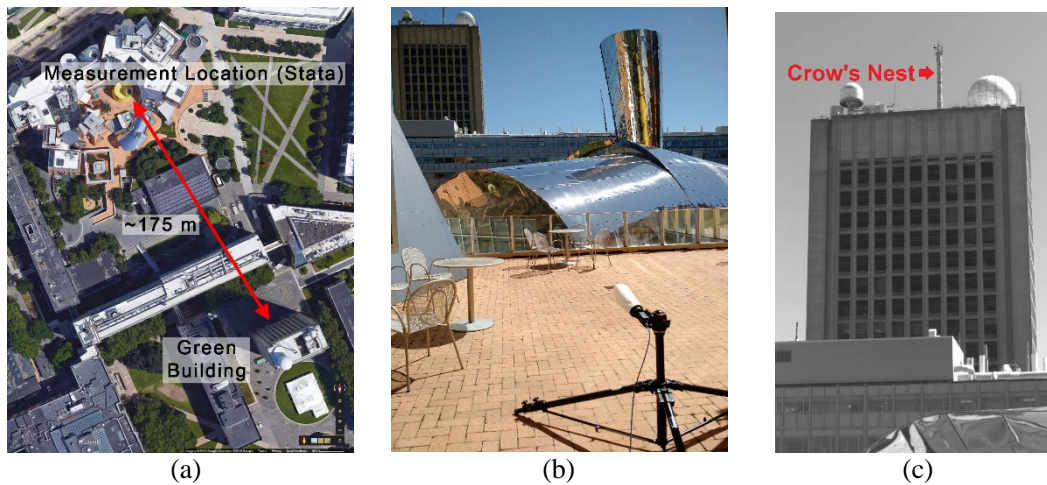


Fig. 7. (a) Satellite video of measurement location, (b) experimental setup for measurement of MIT's Green Building, and (c) recorded video screenshot with crow's nest seen at top

The video was processed using the previously described procedure. Even though the building's motion was too subtle to recover from this distance, vibrations of the crow's nest on top of the building, were strong enough to generate a recoverable signal. Shown in Fig. 8a is the cropped video of the crow's nest, with Fig. 8b showing high contrast pixels with displacements extracted from the video. These displacements were averaged to obtain a single signal for the structure, and an FFT was taken to determine the frequency spectrum shown in Fig. 8c. In the frequency spectrum there is a resonant peak at 2.433 Hz which suggests that the crow's nest atop the building has a resonant mode at that frequency. The amplitude of the resonant peak from the crow's nest works out to be 0.21 mm. The noise floor for the measurement is approximately 0.07 mm, which gives a signal to noise ratio of 3. For these measurement parameters, any structural motion above the noise floor would be measurable. Improvements can be made by using a more telephoto zoom lens in addition to a higher resolution camera.

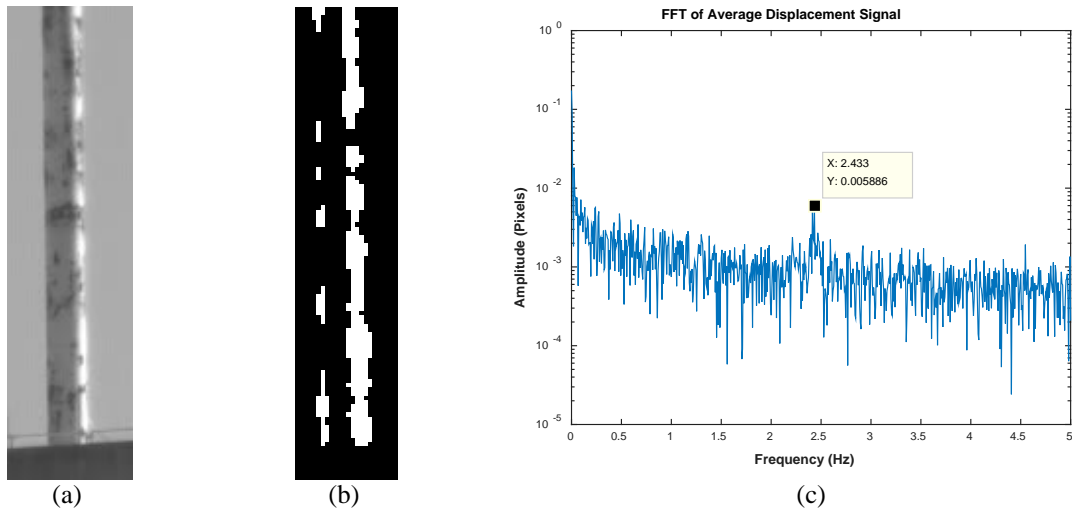


Fig. 8. (a) Screenshot of video cropped to crow's nest, (b) pixel mask of valid displacements, and (c) results from camera measurement of the Green Building crow's nest

As verification for the resonant frequency of the crow's nest measured by the camera, a laser vibrometer was used to measure the frequency response during a day with similar weather conditions. Fig. 9a shows the measurement setup on the roof of the Green Building, measuring the crow's nest vibrations from close range. A measurement was also made of the ground next to the crow's nest as a reference so that any vibrations of the tripod itself could be discounted. The laser vibrometer measurement is shown in Fig. 9b with the crow's nest signal shown in blue and the reference signal in red. Two potential resonant peaks are seen in the crow's nest measurement, however the peak at around 1 Hz is also seen in the reference measurement and thus not a resonance of the crow's nest. The other peak occurs at 2.474 Hz which is similar to the 2.433 Hz measured by the camera; this lends credibility to the camera measurement.

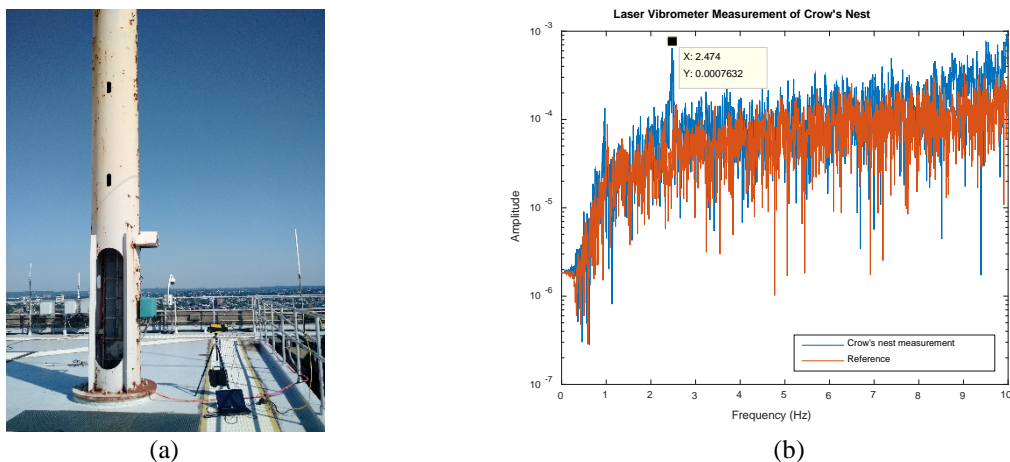


Fig. 9. Laser vibrometer verification of MIT Green Building crow's nest resonant frequency with (a) measurement setup and (b) result with static reference for comparison

2.5 Granite Fracture Test

This application involves a high-speed video of a piece of granite with manufactured defects undergoing a standard compression test, where the quantitative values of interest are the displacements and directions of motion at different locations in the specimen. This is in contrast to many of the previous applications where the measurement is aimed at determining the frequency spectrum of the object in the video. The video was recorded at 14,000 fps with

a resolution of 640 x 792 pixels over 4828 frames of video; a screenshot of the video is shown in Fig. 10a.

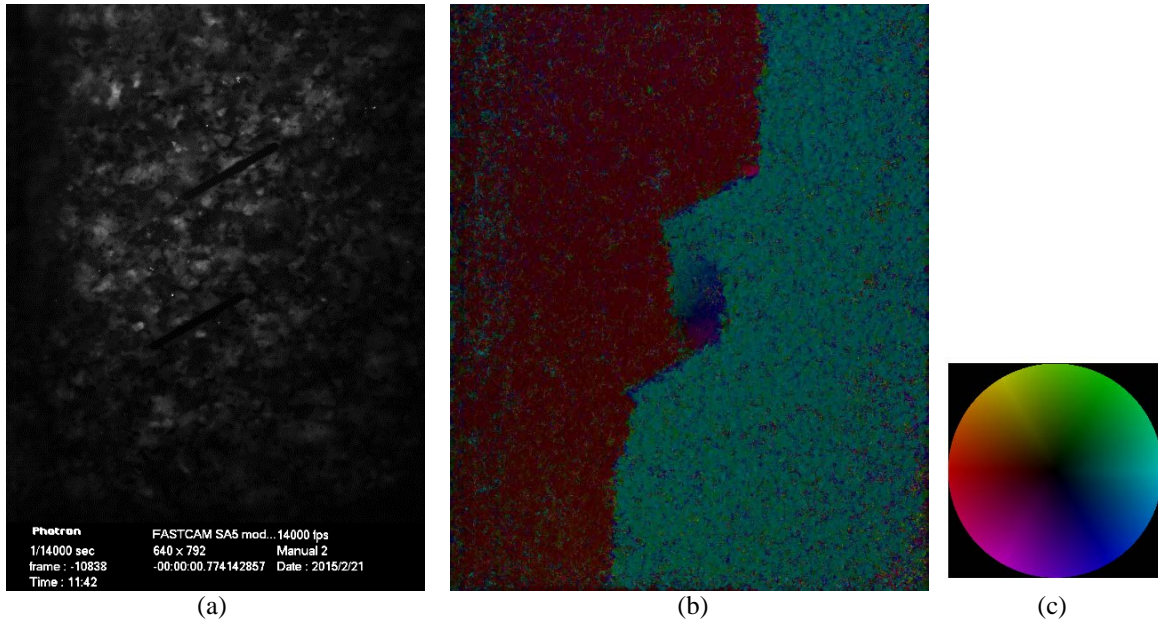


Fig. 10. (a) Screenshot of source video (brightened for visibility), (b) frame from generated video visualizing motion, and (c) color representation of direction and magnitude of motion, e.g. red for left, teal for right

The final 1329 frames were processed to determine the displacement at every pixel in the frame, with the granite providing sufficient visual texture to properly measure the displacements. To visualize these displacements, they were converted into a video representation with the hue (color) signifying the direction, and the brightness representing the magnitude of the displacement as shown in Fig. 10c. To fit within the range of values for a video, the displacements were normalized and the fourth root of the values was taken to magnify smaller values. A frame from the visualization just after fracture occurs is shown in Fig. 10b. A sample horizontal displacement measured for the pixel located at (213,115) in the video is shown in Fig. 11a on a screenshot of the video, with Fig. 11b clearly showing the large amount of movement at fracture, and relaxation afterwards in units of pixels.

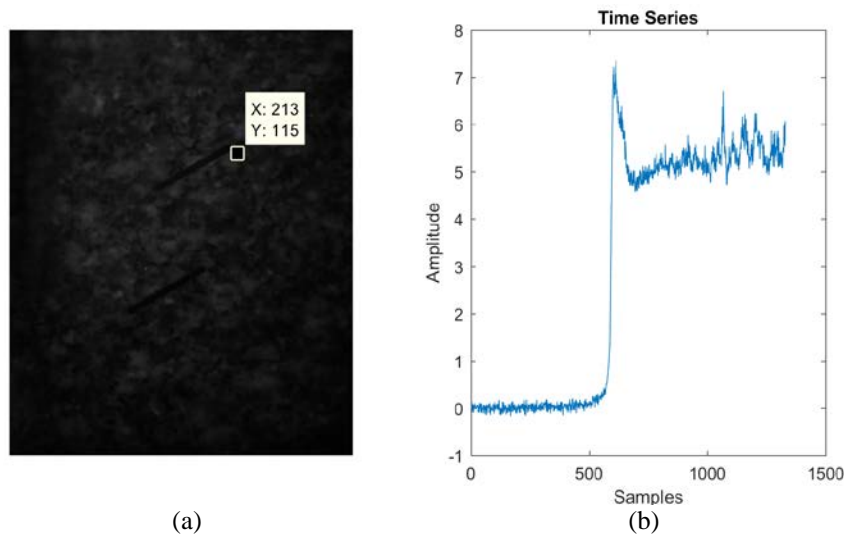


Fig. 11. (a) Screenshot of source video with pixel selected and (b) measured displacement time series of the selected pixel, amplitude in units of pixels

3. Discussion

In this section we discuss the capabilities, limitations, and potential future work and applications for camera measurement of structural motions and vibrations.

3.1 Capabilities and Limitations

Motion magnification makes it very simple to visualize the vibrations and operational mode shapes of an object, something that may be cumbersome with traditional sensors if there isn't a preexisting model for the object. The camera can also be used in a non-contact way to measure the displacements of structures without altering the dynamics of the system. The camera also measures displacements which may otherwise be difficult to measure, since most current vibration sensors measure acceleration, velocity, or strain.

There are several assumptions and limitations of the current measurement and processing methodology. Small motion is a necessary requirement of this processing. This generally means that the motion should be less than one pixel in amplitude, however the video can be downsampled to change the scale of the motion to make it small. A stationary camera is ideal, as any motion of the camera will cause apparent motion of the objects in the scene. Typically this means that video needs to be filmed with a sturdy tripod, however even so a mechanically noisy environment may also corrupt the result. Currently video filmed from a moving platform is generally too unstable for motion magnification or displacement measurement. A limitation is that motion signals are only accurate in areas with good local contrast, as it is not possible to determine the motion of textureless regions in a video. This means displacements are most easily found on the edges of objects with contrast against a background, or with objects painted with a speckle pattern.

3.2 Future Work

Future work involves improving the processing procedure to run even faster so that higher frame rate and larger videos can be handled in real-time. For handling camera motion, work is being done to determine and subtract the motion from known non-moving reference objects in the video, and also use sensors on the camera itself to measure and correct for its motion. For even longer distance measurements atmospheric turbulence is an issue that can also cause apparent motion in the video. Work will need to be done to remove or correct those effects, potentially with a stereo or multiple camera setup.

4. Conclusion

Motion magnification refers to a collection of computer vision algorithms that are used to qualitatively visualize and quantitatively measure small motions of objects in videos. We summarized the video processing procedure and gave example applications for a lightweight beam, rotating machinery, measurement of civil infrastructure both from an indoor security video and an outdoor long range measurement, and a material compression test. The unique capabilities of a camera for non-contact measurement of displacement with potentially low cost equipment opens up a wide range of opportunities in NDT and SHM applications.

Acknowledgements

This research was supported by Shell and we thank chief scientists Sergio Kapusta and Dirk Smit for their support. We also thank Maryse Vachon for providing the lightweight beams,

Sean Kinderman for access and assisting with the MIT facilities measurement, and Stephen Morgan and Bing Li for the granite video.

References

- [1] Wadhwa, N., Rubinstein, M., Durand, F., & Freeman, W. T. (2013). "Phase-based video motion processing." *ACM Transactions on Graphics (TOG)*, 32(4), 80.
- [2] Wadhwa, N., Rubinstein, M., Durand, F., & Freeman, W. T. (2014, May). "Riesz pyramids for fast phase-based video magnification." In *Computational Photography (ICCP), 2014 IEEE International Conference on* (pp. 1-10). IEEE.
- [3] Chen, J. G., Wadhwa, N., Cha, Y. J., Durand, F., Freeman, W. T., & Buyukozturk, O. (2015). "Modal identification of simple structures with high-speed video using motion magnification." *Journal of Sound and Vibration*, 345, 58-71.
- [4] Chen, J. G., Wadhwa, N., Durand, F., Freeman, W. T., & Buyukozturk, O. (2015). "Developments with Motion Magnification for Structural Modal Identification Through Camera Video." In *Dynamics of Civil Structures*, Volume 2 (pp. 49-57). Springer International Publishing.
- [5] Davis, A., Bouman, K. L., Chen, J. G., Rubinstein, M., Durand, F., & Freeman, W. T. (2015, June). "Visual vibrometry: Estimating material properties from small motions in video." In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on* (pp. 5335-5343). IEEE.
- [6] Chen, J. G., Davis, A., Wadhwa, N., Durand, F., Freeman, W. T., & Buyukozturk, O. (2015). "Video camera-based vibration measurement for Condition Assessment of Civil Infrastructure", *NDT-CE International Symposium Non-Destructive Testing in Civil Engineering*, Berlin, Germany, Sept. 15-17, 2015.
- [7] Simoncelli, E. P., & Freeman, W. T. (1995, October). "The steerable pyramid: A flexible architecture for multi-scale derivative computation." In *ICIP* (p. 3444). IEEE.
- [8] Fleet, D. J., & Jepson, A. D. (1990). "Computation of component image velocity from local phase information." *International journal of computer vision*, 5(1), 77-104.
- [9] Gautama, T., & Van Hulle, M. M. (2002). "A phase-based approach to the estimation of the optical flow field using spatial filtering." *Neural Networks, IEEE Transactions on*, 13(5), 1127-1136.
- [10] YouTube video. (2015, April 15). <https://youtu.be/LA-ZcEWCOIc>