Title: **Structural Health Monitoring from the Window Seat of a Passenger Airplane**

Authors : Abe Davis[1,2]
Justin G. Chen[3]
Oral Büyüköztürk[3]
Fredo Durand[2]
Doug L. James[1]

ABSTRACT

Recent advances in computer vision and graphics have shown that regular, monocular cameras and video (e.g. cell phone cameras and Digital SLRs) can be used to identify the resonant frequencies of structures, and even image visually subtle operational deflection shapes. This paper is offered as a teaser for that work, focusing specifically on an example that may be of interest to people in the structural health monitoring (SHM) community. The discussion is high-level, and presented in an intentionally casual tone (much of the paper presents an anecdote – about recovering the operational deflection shape of an airplane wing using a cell phone and a dish sponge – using the first person). Our hope is to make this text as accessible and painless to read as possible, with hopes of introducing readers from different engineering disciplines to our related work in computer vision and graphics.

INTRODUCTION

Cameras have become one of the most ubiquitous technologies of our time. Massive demand, driven largely by consumer photography, has led to rapid advances and cost-reduction in imaging technology. As these improvements outpace the development of other technologies, visual solutions to many problems in engineering become increasingly appealing. This paper is offered as an example to illustrate how low cost camera technology could play a role in the future of structural health monitoring (SHM).

The vibrations of structures are most commonly measured with contact sensors like accelerometers that require physical access to a structure and can be labor intensive to put in place. Recently, there has been much interest and development in the use of computer vision based methods for vibration measurement and SHM [1–7]. Cameras have the

Abe Davis, Doug L. James. Stanford University, Department of Computer Science, Stanford University, Stanford, CA 94305, U.S.A..
Abe Davis, Fredo Durand. CSAIL, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, U.S.A.
Justin G. Chen, Oral Buyukozturk, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, U.S.A.

potential to offer an inexpensive, easy to setup, and accurate way to characterize the vibration behavior of structures and track damage over time.

VISUAL VIBRATION ANALYSIS IN VISION AND GRAPHICS

Our work builds on several recent works in computer vision and graphics that deal with recovering, analyzing, and visualizing very small motion in video [8–15]. We begin with a high-level overview of this work and how it relates to our own. For a more detailed summary, refer to [15].

Eulerian Motion Estimation in Video

Like prior work, we use an Eulerian approach to estimating motion in video, based on the complex steerable pyramid [16, 17]. The basic idea behind this approach is that motion can be measured by examining how image gradients (i.e. edges) shift locally around each point in an image. We use this to recover local signals describing the motion around every pixel in an image, across all frames of a video. For the purposes of this paper, we consider separate signals describing the motion at each pixel along different spatial orientations within a video (e.g. one signal describing motion across columns of pixels in a video, and another describing motion across rows). The signal $u(\theta, x, y, t)$ then describes displacement of pixel x, y in the direction θ at time t. For convenience, we will also use the shortened notation $u_\ell(t)$ to indicate local motion signals for some location and orientation θ, x, and y.

In practice, the local motion signals $u_\ell(t)$ are always weighted by a measure of local image contrast, which we will not describe here. This is done to place higher confidence on parts of a video with stronger texture, where motion can be more precisely estimated. Intuitively, if a completely white object were moving in front of a completely white background, we would see no evidence of this motion in a video – we would simply see white on white. Weighing local motion signals by local image contrast ensures that such ambiguous cases won't introduce large amounts of noise into calculations. For a discussion on weighing local motion, refer to [15].

Video Motion Spectra

By taking the Fourier transform of local motion signals $F(u_\ell(t))$, we can examine the spectra of local motions in a video. Recall that modal frequencies ω_i are global properties of an object. This means that the power spectra of all local motion signals $u_\ell(t)$ that measure the same object should have spikes at the same resonant frequencies. This holds even if two points on the object move with opposite phase or in different directions at a given frequency. This observation allows us to average across local *video motion spectra* $U_\ell(\omega) = F(u_\ell(t))$ to find mode frequencies ω_i by looking for peaks in a single global spectrum. To find resonant frequencies, we want local motions with opposite phase to add constructively into our global signal. To accomplish this, we simply add the power spectra of local motion signals to calculate a global power spectrum $P(\omega)$:

$$P(\omega) = \sum_\ell \mid U_\ell(\omega) \mid^2 = \sum_\ell \mid F(u_\ell(t)) \mid^2 \tag{1}$$

Modal Imaging

The global power spectrum $P(\omega)$ will serve as our main tool for identifying the frequencies of operational deflection modes in video. We can use this signal in much the same way that we would use a power spectrum recovered from any other type of vibration sensor. However, the spatial resolution of video also offers us use of the local video motion spectra $U_\ell(\omega)$. While the modal frequencies ω_i are global properties of an object, which we can assume to be shared by all local video motion spectra, local vibrations across the object are *scaled* according to mode shapes ϕ_i. In video, this means that the magnitude of motion from each vibration mode should be scaled by a projection of the corresponding shape ϕ_i. We can estimate this projected shape by computing what we call a *modal image*, which is an image $\mathcal{U}_{\omega_i}(x, y)$ of temporal Fourier coefficients corresponding to motion at a common frequency, or set of frequencies associated with a common mode.

In the simplest case, our modal images are frequency slices of a spectral volume $U(x, y, \omega) = F_t(u(x, y, t))$, where F_t denotes a Fourier transform along the time dimension only. In practice, we weigh and filter our local motion signals before taking their Fourier transform. For details on how that is done, refer to [15].

PREVIOUS APPLICATIONS

The methods described above were developed over a number of different projects, addressing a number of different applications. Here we summarize those applications, with references, for curious readers:

The Visual Microphone [12] This work showed that vibrations caused by sound could be extracted from video to recover audio, effectively turning visible objects to visual microphones from a distance. An earlier version of video motion spectra were used in this work in an attempt to equalize recovered sounds, correcting for the resonance of objects being measured.

Visual Vibrometry [13, 18] This work used video motion spectra to identify the resonant frequencies of objects in video, in order to identify material properties.

Interactive Dynamic Video [14] This work introduced an interface for selecting and visualizing modal images from a global video motion power spectrum, and used selected modal images as a basis for the interactive simulation of objects in video.

MODAL IMAGING IN THE WILD

Author's Note: In this section, I (Abe Davis) use the first person to present an experiment through anecdote. This is done to reflect the spur-of-the-moment nature of the experiment, and how our work has the potential to enable SHM in casual settings, with common, everyday equipment.

Modal imaging can be a powerful tool for analyzing vibrations in real world setting. Cameras offer a number of significant advantages over other more traditional vibration sensors. Because cameras are passive and offer significant spatial resolution, it is often possible to simply point a camera at an object and analyze its vibrations with little setup or advanced planning. Figure 1 shows an anecdotal but compelling example.

In the summer of 2015 I took a flight from Boston to San Francisco to present our paper "Long Distance Video Camera Measurements Of Structures" [19] at the 10th International Workshop on Structural Health Monitoring (IWSHM 2015), held at Stanford. As third author I had not expected to present the paper, but inherited the task when the first two authors became unavailable.

Noting that several other papers at the conference focused on SHM of aircraft using modal analysis, I decided to run an experiment from the window seat of my flight to the conference. Using a sponge, I propped my cell phone up against the window (Figure 1 top left), pointed at the airplane's wing, and recorded about two minutes of 30fps video (Figure 1 bottom left).

Upon arriving at my hotel for the workshop, I selected 102 seconds of clean video (cropping portions where I had to adjust the camera at the beginning and end) and ran projection-only video stabilization on the remaining footage using Adobe After Effects CS6 (meaning that I only stabilized frame-to-frame motion that was uniform across the entire image plane). I then uploaded the stabilized video to our server at MIT and ran our modal imaging code. The right half of Figure 1 shows our mode selection interface (which lets users visualize modal images at selected frequencies) and a selected mode at 2.5245Hz. I predicted that this mode was the dominant bending mode of the airplane wing, and reported the result to Justin G. Chen, my coauthor from the Civil and Environmental Engineering Department at MIT.

Justin was able to find the model aircraft that I had taken by looking up the flight number. He was then able to find a reference for the dominant flex mode of that aircraft, which was reported as 2.584 Hz [20]. This is less than 3% off from the frequency we predicted from video. Considering the effects of temperature and fuel in the wing, and that the video was taken on the fly (literally) with no special equipment (other than a common dishwashing sponge) this is a rather impressive result that illustrates potential for casual and ubiquitous SHM.

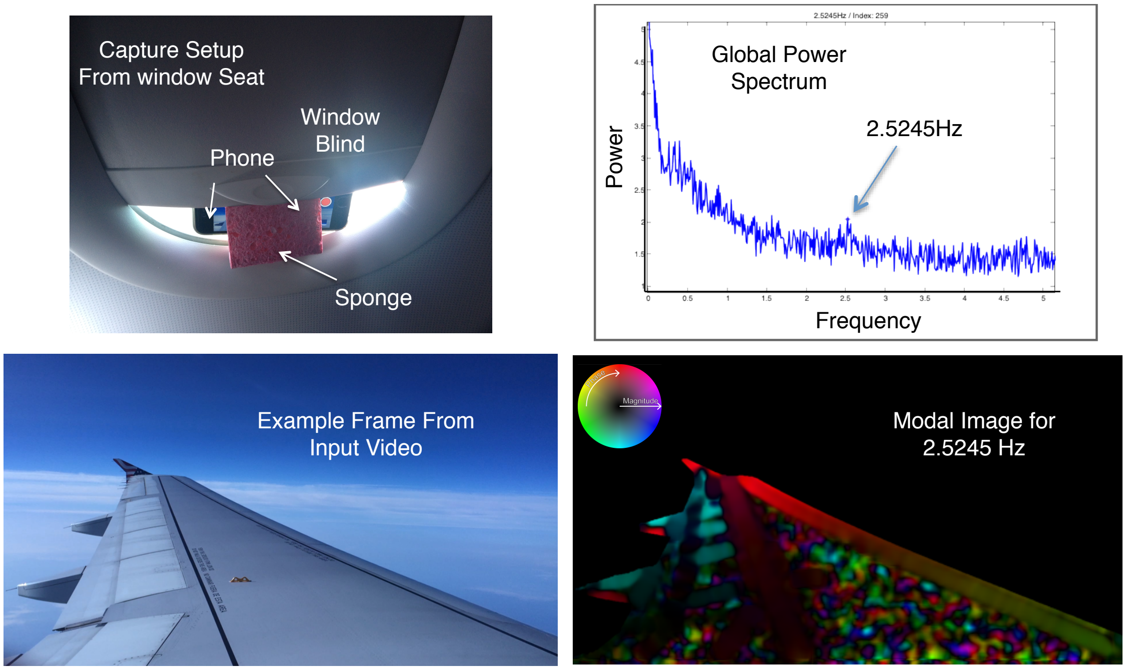

Figure 1. On a flight from Boston to San Francisco to present our paper [19] at the International Workshop on Structural Health Monitoring (IWSHM), I filmed the airplane wing from a window seat in the cabin using my cell phone. To record the video, I wedged my phone under the window blind with a common household sponge (top left). On the bottom left we see a frame from the captured video. On the top right we see the global power spectrum recovered from the video. On the bottom left we see the modal image for a predicted flex mode at 2.5245 Hz, which is less than 3% off from the reported value of 2.584 Hz [20]. The modal image is for the x dimension of the video. The opposite blue/red phase relationship on parts of the wing are likely the result of rotation relative to the optical axis of the camera.

CONCLUDING REMARKS

This anecdote provides an example for how accurate vibration measurements can be made with a smartphone. Cameras can range from extremely expensive high-speed laboratory units to inexpensive cell phone and web cameras. Smartphones such as the iPhone 7 plus and the Google Pixel can already record at resolutions (4K) and frequencies (240 fps) with sufficient quality to measure the vibrations of most large civil and aerospace structures and machinery. There are some issues to be worked out (such as the effects of camera motion and video compression), but consumer cameras have proven to be useful for video-based methods for structural health monitoring, and the ubiquity of these devices offers exciting new possibilities for the field.

REFERENCES

1. Lee, J. J. and M. Shinozuka. 2006. “A vision-based system for remote sensing of bridge displacement,” *Ndt & E International*, 39(5):425–431.
2. Schumacher, T. and A. Shariati. 2013. “Monitoring of structures and mechanical systems using virtual visual sensors for video analysis: Fundamental concept and proof of feasibility,” *Sensors*, 13(12):16551–16564.

3. Yang, Y., C. Dorn, T. Mancini, Z. Talken, G. Kenyon, C. Farrar, and D. Mascareñas. 2017. "Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification," *Mechanical Systems and Signal Processing*, 85:567–590.
4. Bao, Y., P. Seshadri, and S. Mahadevan. 2017. "Motion Magnication For Quantifying Aeroelastic Modes From High-Speed Videos," in *58th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, p. 0869.
5. Sarrafi, A., P. Poozesh, C. Niezrecki, and Z. Mao. 2017. "Mode extraction on wind turbine blades via phase-based video motion estimation," in *SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring*, International Society for Optics and Photonics, pp. 101710E–101710E.
6. Chen, J. G., N. Wadhwa, Y.-J. Cha, F. Durand, W. T. Freeman, and O. Buyukozturk. 2015. "Modal identification of simple structures with high-speed video using motion magnification," *Journal of Sound and Vibration*, 345:58–71.
7. Chen, J. G., A. Davis, N. Wadhwa, F. Durand, W. T. Freeman, and O. Büyüköztürk. 2017. "Video Camera–Based Vibration Measurement for Civil Infrastructure Applications," *Journal of Infrastructure Systems*, 23(3).
8. Wu, H.-Y., M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman. 2012. "Eulerian video magnification for revealing subtle changes in the world," *ACM Transactions on Graphics (TOG)*, 31(4):65.
9. Wadhwa, N., M. Rubinstein, F. Durand, and W. T. Freeman. 2013. "Phase-based video motion processing," *ACM Transactions on Graphics (TOG)*, 32(4):80.
10. Wadhwa, N., M. Rubinstein, F. Durand, and W. T. Freeman. 2014. "Riesz Pyramid for Fast Phase-Based Video Magnification," in *Computational Photography (ICCP), 2014 IEEE International Conference on*, IEEE.
11. Rubinstein, M. 2014. *Analysis and Visualization of Temporal Variations in Video*, Ph.D. thesis, Massachusetts Institute of Technology.
12. Davis, A., M. Rubinstein, N. Wadhwa, G. J. Mysore, F. Durand, and W. T. Freeman. 2014. "The Visual Microphone: Passive Recovery of Sound from Video," *ACM Trans. Graph.*, 33(4):79:1–79:10, ISSN 0730-0301, doi:10.1145/2601097.2601119.
13. Davis, A., K. L. Bouman, J. G. Chen, M. Rubinstein, F. Durand, and W. T. Freeman. 2015. "Visual Vibrometry: Estimating Material Properties From Small Motion in Video," .
14. Davis, A., J. G. Chen, and F. Durand. 2015. "Image-space Modal Bases for Plausible Manipulation of Objects in Video," *ACM Trans. Graph.*, 34(6):239:1–239:7, ISSN 0730-0301, doi:10.1145/2816795.2818095.
15. Davis, A. 2016. Ph.D. thesis, Massachusetts Institute of Technology.
16. Simoncelli, E. P., W. T. Freeman, E. H. Adelson, and D. J. Heeger. 1992. "Shiftable multiscale transforms," *IEEE Trans. Info. Theory*, 2(38):587–607.
17. Portilla, J. and E. P. Simoncelli. 2000. "A Parametric Texture Model Based on Joint Statistics of Complex Wavelet Coefficients," *Int. J. Comput. Vision*, 40(1):49–70, ISSN 0920-5691, doi:10.1023/A:1026553619983.
18. Davis, A., K. L. Bouman, J. G. Chen, M. Rubinstein, O. Bykztrk, F. Durand, and W. T. Freeman. 2017. "Visual Vibrometry: Estimating Material Properties from Small Motions in Video," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4):732–745, ISSN 0162-8828, doi:10.1109/TPAMI.2016.2622271.
19. Chen, J. G., N. Wadhwa, A. Davis, F. F. Durand, T. William, and O. BUYUKOZTURK. 2015. "Long Distance Video Camera Measurements Of Structures," *10th International Workshop on Structural Health Monitoring (IWSHM 2015)*.
20. BUCHARLES, A., H. CASSAN, and J. ROUBERTIER. 1990. *Advanced parameter identi-*

fication techniques for near real time flight flutter test analysis, American Institute of Aeronautics and Astronautics, doi:doi:10.2514/6.1990-1275.