# Image Space Modal Analysis and Simulation of Object Motion

Brad Huang
brad0309@stanford.edu

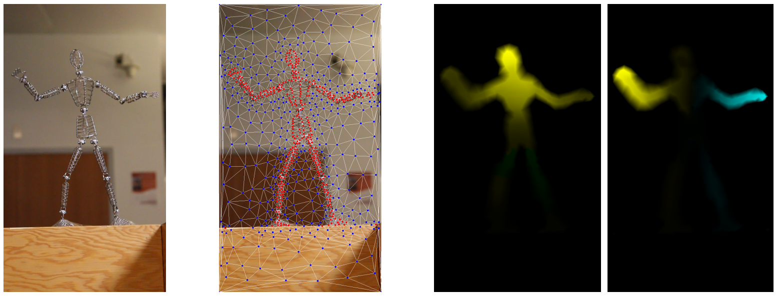

**Figure 1:** *Wireman in different parts of the pipeline, including original frame, feature points and triangulation, and X and Y modal shapes.*

## Abstract

For the final project, I will work on improving my previous project on interactive dynamic video to combine and leverage advantages of different approaches. The interactive dynamic video is a research project from Abe Davis [Davis et al. 2015a], where they extracted image space modal bases from videos of vibrating objects, and used the knowledge to simulate interaction with the objects. The method they used to extract modal bases is considered Eulerian - optical flow of each pixel in the first frame is computed against every subsequent frames, and the flow vectors are then used to perform modal analysis.

The extension that we are working on, is to instead use a Lagrangian approach, where we extract feature points of the objects from the first frame, triangulate the frame with the feature points, track and perform modal analysis on those feature points, and interpolate between them when simulating to move the facets.

The benefits of the Lagrangian approach focus on the ability to correct for global motion stemming from camera movements, and to separate different components of the object to avoid false coupling of vibrational modes. However, the approach also suffers from artifacts from the interpolation, which mainly appear in the boundary of the object (you can look at my previous video to see how it looks like). In addition, the lack of depth information leads to some undesired modal basis from motion blur and parallax.

Abe and I have discussed several ways to reduce artifacts from the Lagrangian approach. The most apparent solution would be including depth information, so that foreground and background can be easily segregated, and including depth in the analysis and simulation would create more realistic interactions. However, owing to the lack of high speed and high precision depth camera out there, we decided to try using the Eulerian modes as a way to regularize the motion of interpolated facets. In essence, we would be looking into how to best combine the information from the two approaches to preserve both advantages.

## 1 Background Information

I would refer to my previous project report for Prof. James' class CS348C in Fall 2016 (attached with this proposal) for background information on how the Lagrangian analysis method works, and how the entire pipeline is pieced together. Figure 1 shows the intermediate results from the project, with the reference frame of the wireman video, the triangulation of the reference frame with feature points and fixed points, and visualization of one modal basis that exhibit the behavior of moving the head in response to moving the hands. As described in the paper of visual vibrometry [Davis et al. 2015b], the visualization of the modal basis uses HSV color space, with the phase as the hue, and the magnitude as the value. As a result, the x and y modal shapes show that when the left hand moves up, the right hand moves down since its y mode is in the opposite phase, while the head moves right since its x mode is in the same phase.

## 2 Related Work

**Observing Vibration Modes:** Directly observing modal shapes from videos has been a topic discussed widely in the optical and mechanical engineering society [Helfrick et al. 2011; Bianchi 2014]. In addition, there has also been research in computer graphics and vision that seeks to identify [Wadhwa et al. 2014; Chen et al. 2015] vibrational modes for the purpose of motion magnification and visualization. Our approach, as extensively investigated in Abe's thesis [Davis 2016], focuses on modal shapes that gain knowledge of object properties in the scene and that are best for simulation of object responses.

**Motion Magnification:** Recent publications in motion magnification of object motion [Wadhwa et al. 2014; Chen et al. 2015] have demonstrated the ability to extract and magnify motion from observations. Interactive dynamic video, as in [Davis et al. 2015a] and our project, is different since by learning about vibration modes from videos, we are able to synthesize motion in response to different forces that were not present in the videos to make the final result interactive.

## 3 Project Overview

This project would focus on using the modal shapes obtained from the Eulerian and Lagrangian approaches to achieve better simulation results. I believe there are two main ways we can do it:

- Analysis: the modal shapes can be easily visualized as images, as shown above in Figure 1. For videos captured with stationary camera and contain less variations in vibration modes as in the wireman video, the main difference between the modal shapes produced from the two approaches is how sharp the edges are around the vibrating objects. For objects that are complicated and may have dense variations in the structure of the modal shapes, the Lagrangian approach could introduce undesirable structures at times, although it does reduce the noise that may come from small spatial variations. As a result, one way to combine the results would be using the Eulerian modal shapes as a prior to regularize the modal shape of the Lagrangian modal shapes.

- Simulation: at simulation time, it is also good to consider using the Eulerian modal shapes to accommodate for effects like motion blur and parallax. If you compare the products of my video with Abe's video, the Eulerian approach gave some hints to motion blur and parallax. The reason is that with the Eulerian approach, the effects are much more visible when they cover only a small area of the object. For example, you only see motion blur with the outer boundary of the wireman's chest, and parallax with some parts of the bush. To discover these details, the Lagrangian approach would need to use very dense feature points around those areas to avoid disregarding them with interpolation. One way to fix this is to avoid using naive interpolation for non-feature point areas in the triangulation, but instead uses the information from the Eulerian modal shape to correct the resulting motion. This can take the form of adjusting texture coordinates in the shader, or computing a weighted interpolation scheme based on the Eulerian modal shapes beforehand.

## 4 Milestone, Timeline and Goals

- **Merge analysis pipelines (2/20):** Construct an easy to use interface and code base that can generate both types of analysis results (dense and sparse optical flow).
- **Simulation platform (2/27):** Refine the simulation platform so that we can feed arbitrary mode shapes into it and perform simulation and visual debugging on it.
- **Modal shape combination (3/6):** Experiment with different ways of combining the modal shapes of the two approaches, and visualize how effective they are using the simulation platform. Quantitatively examine the results can be done by calculating MSEs between results from the hybrid approach and the individual approaches.
- **Simulation refinement (3/13):** Experiment with ways to use the Eulerian modal shapes in the simulation of the interactions.
- **Wrap up (3/17):** Wrap up by generating examples and results that illustrate the strength of the hybrid approach, and prepare for presentation and posters.

## References

ANONYMOUS, 1976. Planes of the head. `http://www.planesofthehead.com/`.

BIANCHI, S. 2014. Vibration detection by observation of speckle patterns. *Optical Society of America*.

CHEN, J. G., WADHWA, N., CHA, Y., DURAND, F., FREEMAN, W. T., AND BUYUKOZTURK, O. 2015. Modal identification of simple structures with high-speed video using motion magnification. *Journal of Sound and Vibration*.

DAVIS, A., CHEN, J., AND DURAND, F. 2015. Image-space modal bases for plausible manipulation of objects in video. In *Proceedings of SIGGRAPH Asia 2015*, ACM Transactions on Graphics (TOG), New York, vol. 34, ACM.

DAVIS, A., DOUMAN, K. L., CHEN, J. G., RUBINSTEIN, M., AND DURAND, F. 2015. Visual vibrometry: Estimating material properties from small motion in video. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

DAVIS, A. 2016. *Visual Vibration Analysis*. PhD thesis, MIT CSAIL.

HELFRICK, M., NIEZRECKI, C., AVITABILE, P., AND SCHMIDT, T. 2011. 3d digital image correlation methods for full-field vibration measurement. *Mechanical Systems and Signal Processing*.

WADHWA, N., RUBINSTEIN, M., DURAND, F., AND FREEMAN, W. T. 2014. Riesz pyramids for fast phase-based video magnification. *2014 IEEE International Conference on Computational Photography (ICCP)*.