Visual Vibrometry: Estimating Material Properties from Small Motions in Video

Abe Davis*[1], Katherine L. Bouman*[1], Justin G. Chen[1], Michael Rubinstein[2], Frédo Durand[1], William T. Freeman[1,2]
[1]Massachusetts Institute of Technology [2]Google Research * Denotes Joint First Author

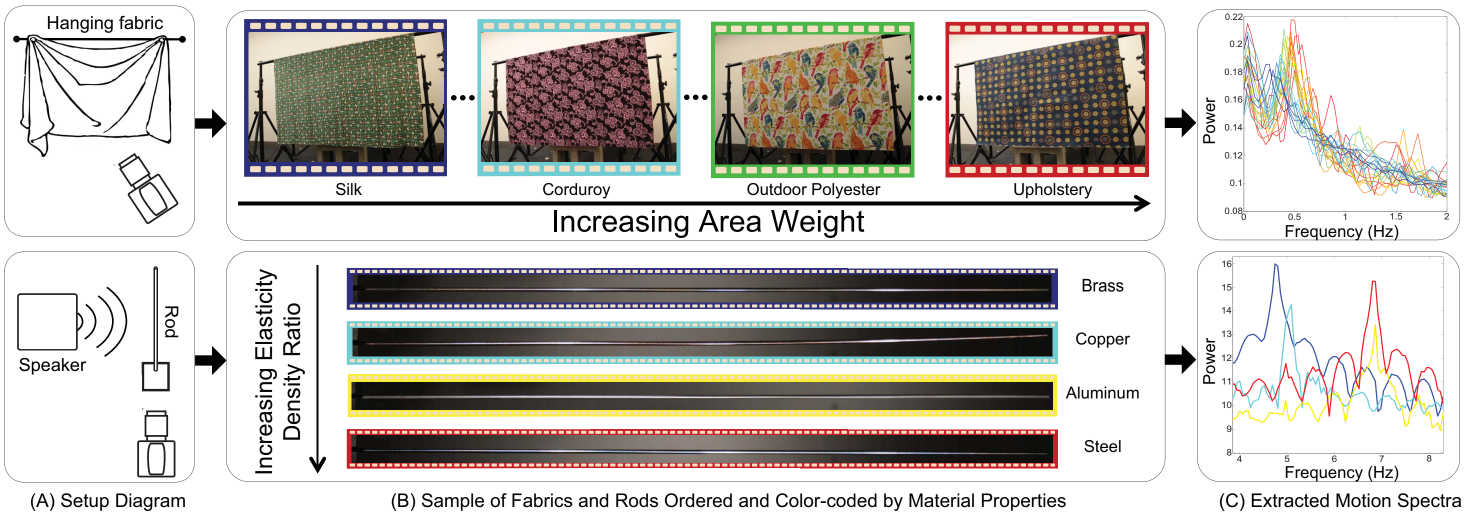

Figure 1: We present a method for estimating material properties of an object by examining small motions in video. (A) We record video of different fabrics and clamped rods exposed to small forces such as sound or natural air currents in a room. (B) We show fabrics (top) color-coded and ordered by area weight, and rods (bottom) similarly ordered by their ratio of elastic modulus to density. (C) Local motion signals are extracted from captured videos and used to compute a temporal power spectrum for each object. These motion spectra contain information that is predictive of each object's material properties. For instance, observe the trends in the spectra for fabrics and rods as they increase in area weight and elasticity/density, resp (blue to red). By examining these spectra, we can make inferences about the material properties of objects.

The estimation of material properties is important for scene understanding, with many applications in vision, robotics, and structural engineering. This work connects fundamentals of vibration mechanics with computer vision techniques in order to infer material properties from small, often imperceptible motion in video. Objects tend to vibrate in a set of preferred modes. The shapes and frequencies of these modes depend on the structure and material properties of an object [6]. Focusing on the case where geometry is known or fixed, we show how information about an object's modes of vibration can be extracted from video and used to make inferences about that object's material properties. We demonstrate our approach by estimating material properties for a variety of rods and fabrics by passively observing their motion in high-speed and regular-framerate video [5].

Understanding a scene involves more than just recognizing object categories or 3D shape. The physical properties of objects, such as the way they move and bend, can be critical for applications that involve assessing or interacting with the world. In the field of non-destructive testing, an object's physical properties are often studied through the measurement of its vibrations using contact sensors or expensive laser vibrometers [2, 7]. In both cases, measurements are often limited to a small set of discrete points. In contrast, we leverage the ubiquity and high spatial resolution of video cameras to extract physical properties from video. In recent work, Chen et al. [3] used videos to quantify the vibration modes of cantilever beams. We extend similar ideas to automatically extract physical properties from video and make inferences about an object's underlying material properties. We are inspired by recent work in computer vision, but seek to bridge the gap with engineering techniques and focus on fundamentals of vibration analysis.

Objects tend to vibrate in a set of preferred modes. These vibrations occur in most materials, but often happen at scales and frequencies outside the range of human visual perception. Bells, for instance, vibrate at distinct audible frequencies when struck. We cannot usually see these vibrations because their amplitudes are too small and their frequencies are too high - but we hear them. Intuitively we know that large bells tend to sound deeper than small ones, or that a bell made of wood will sound muted compared to one made of silver. This is because an object's modes of vibration are closely related to its geometry and material properties. In this work, we show how this connection can be used to estimate the material properties of an object with fixed or known geometry from video.

We connect established theory of modal vibrations to features that can be extracted from video. We then show how these features can be used to estimate the material properties of objects with fixed or known geometry. We demonstrate this on two sets of objects: clamped rods and hanging fabrics. With each set of objects we explore a different method to resolve the ambiguous contribution of structure (geometry) and material properties to an object's vibrations. Our rod experiments accomplish this with careful measurements in a setting that resembles typical engineering applications. Our fabric experiments instead explore the potential of a learning approach with more direct applications in computer vision.

We use small local motions in video to reason about the modes of recorded objects. For each spatial point in a video, we compute the local motion around that point over time [4]. Our analysis relates the spectra of these motion signals to mode shapes A_i and frequencies ω_i. By examining the temporal power spectra (Figure 1(C)), we can estimate the material properties of a previously unseen, fixed or known geometry object.

1 Estimating Properties of Materials with Known Geometry: Rods

In our first experiments we estimate the material properties of various rods by extracting their resonant frequencies from video. The simple geometry of a clamped rod makes it easy to solve for vibration modes analytically as a function of length, diameter, density, and an elastic modulus. While length, diameter, and density can all be measured with a simple ruler and scale, the elastic modulus is usually measured with a tensile test, which requires expensive equipment and usually damages the object being tested. In these experiments we show how this elastic modulus can instead be measured with a speaker and high-speed camera.

Sound from a loudspeaker is used to induce tiny motions in the rods. We then extract these motions and find the rod's resonant frequencies in the motion's power spectrum. Under fixed but unknown geometry, the recovered fundamental frequencies provide a value proportional to $\sqrt{\frac{E}{\rho}}$ [6]. We use

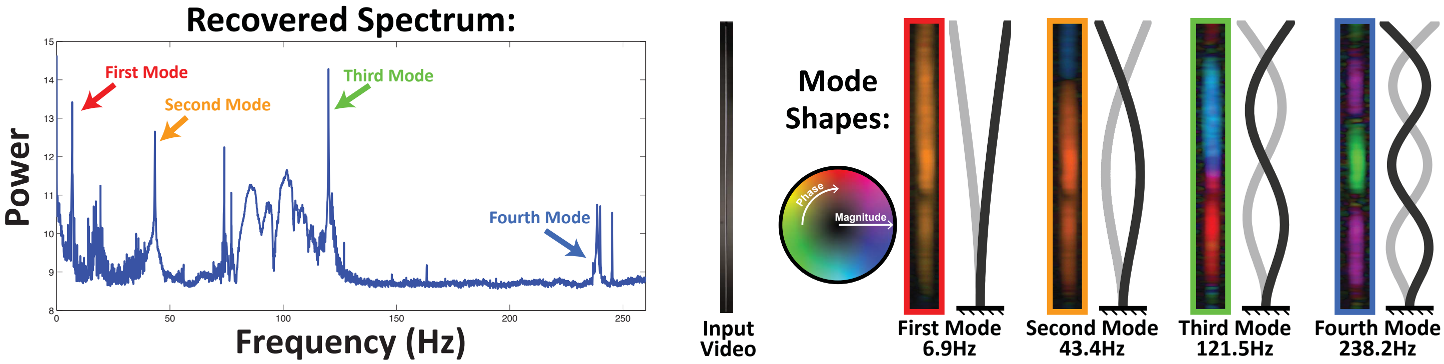

Figure 2: Finding vibration modes of a clamped brass rod: (Left) We recover a motion spectrum from 2.5 kHz video of a 22 inch clamped aluminum rod. Resonant frequencies are labeled. To distinguish resonant frequencies from other spikes in the spectrum, we look for energy at frequencies with ratios derived from the known geometry of the rod. (Middle) A sample frame from the 80×2016 pixel input video. (Right) Visualizations of the first four recovered mode shapes are shown next to the corresponding shapes predicted by theory.

this information along with the lengths and densities measured by a scale and measuring tape to compute the modulus of each rod. Figure 3 shows a plot of Young's moduli reported by the manufacturer against the values estimated using our technique. Error bars are calculated for each moduli by propagating error bounds for length, diameter, and density.

For each rod, we can further verify recovered modes by visualizing the recovered shapes corresponding to estimated resonant frequencies (see Figure 2). In practice we see the predicted shapes of multiple modes in the data recovered for each rod.

2 Learning Properties of Materials with Unknown Geometry: Fabrics

The inference described in the previous section relies on knowing the ratios between resonant frequencies. These ratios are simple to derive in clamped rods, but can be prohibitively difficult to compute in more general structures. As a result, many applications of vibrometry are limited to simple geometries that can be precisely measured (as is the case with rods) or man-made structures (airplanes, buildings, cars, etc) with resonant frequencies that can be derived from detailed CAD models through FEM analysis. The ubiquity and passive nature of video offers the potential to address this limitation by providing sufficient data to learn relationships between motion spectra and the material properties of objects. We have explored that potential by using a learning approach to estimate the material properties of hanging fabrics from video. In this work we have used a dataset of 30 fabrics along with ground truth measurements of stiffness and area weight collected by Bouman, et al. [1].

Motion in the fabrics is either induced using sound from a loudspeaker or purely from ambient forces in the room. After extracting the tiny motion, we learn a regression model that maps the motion spectra to the log of ground truth stiffness or area weight measurements provided in [1].Due to the small number of fabrics in the dataset, we use a leave-one-out method for training and testing. Precisely, all data corresponding to a fabric are removed from training of the regression parameters when predicting the material properties of that fabric. Using this method, we estimate the performance of our model on predicting the material properties of a previously unseen fabric.

Our estimates of material properties are well correlated with the ground truth measurements. Figure 4 contains correlation plots that compare our algorithm's predicted measurements of stiffness and area weight to the log of ground truth measurements when models were trained and tested on videos of fabrics excited by ambient forces and acoustic waves separately. In all cases, even when testing under conditions with different viewpoints and excitation forces from the training data, our estimates outperform the current state of the art algorithm [1] in predicting both stiffness and area weight.

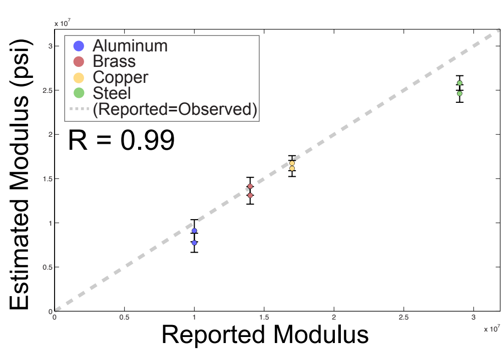

Figure 3: Estimating the elastic modulus of clamped rods: Young's moduli reported by the manufacturer are plotted against values estimated using our technique for aluminum, brass, copper, and steel. Estimated values are close to those reported by the manufacturer.

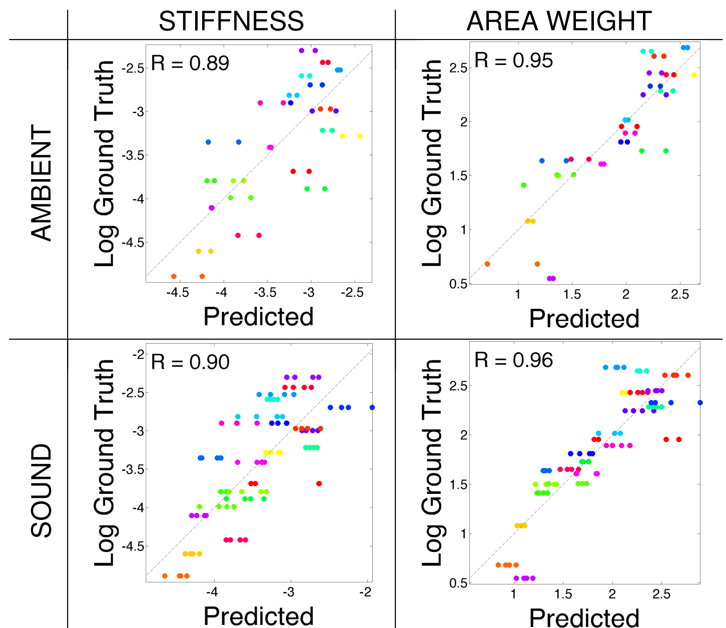

Figure 4: Comparisons between ground truth and model predictions on material properties estimated from videos of fabric excited by ambient forces and acoustic waves. Each circle in the plots represents the estimated properties from a single video. Identical colors correspond to the same fabric. The Pearson product-moment correlation coefficient (R-value) averaged across video samples containing the same fabric is displayed.

[1] Katherine L. Bouman, Bei Xiao, Peter Battaglia, and William T. Freeman. Estimating the material properties of fabric from video. *Computer Vision, IEEE International Conference on*, 0:1984–1991, 2013. ISSN 1550-5499. doi: http://doi.ieeecomputersociety.org/10.1109/ICCV.2013.455.

[2] P. Castellini, M. Martarelli, and E.P. Tomasini. Review. *Mechanical Systems and Signal Processing*, 20(6):1265–1285, 2006. doi: 10.1016/j.ymssp.2005.11.015.

[3] Justin G Chen, Neal Wadhwa, Young-Jin Cha, Frédo Durand, William T Freeman, and Oral Buyukozturk. Structural modal identification through high speed camera video: Motion magnification. In *Topics in Modal Analysis I, Volume 7*, pages 191–197. Springer, 2014.

[4] Abe Davis, Michael Rubinstein, Neal Wadhwa, Gautham J. Mysore, Frédo Durand, and William T. Freeman. The visual microphone: Passive recovery of sound from video. *SIGGRAPH*, 33(4), July 2014.

[5] Abe Davis, Katherine L. Bouman, Justin G. Chen, Michael Rubinstein, Fredo Durand, and William T. Freeman. Visual vibrometry: Estimating material properties from small motion in video. June 2015.

[6] Ahmed A Shabana. *Theory of vibration*, volume 2. Springer, 1991.

[7] P.J. Shull. *Nondestructive evaluation: theory, techniques, and applications*, volume 142. CRC, 2002.