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]
[1]Massachusetts Institute of Technology [2]Google Research

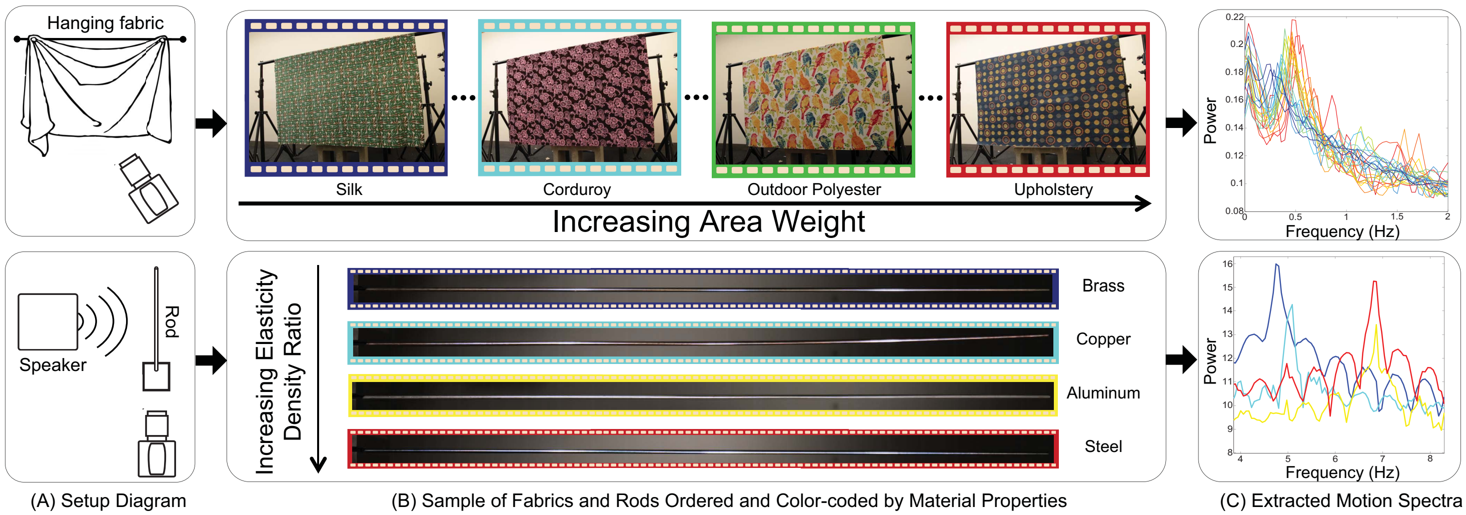

Figure 1: We present a method for estimating material properties of an object by examining small motions in video. (A) We record video of different fabrics and clamped rods exposed to small forces such as sound or natural air currents in a room. (B) We show fabrics (top) color-coded and ordered by area 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, respectively (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 paper 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. 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.

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. 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. These physical properties are then used to make inferences about the 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 paper, we show how this connection can be used to estimate the material properties of an object with fixed or known geometry from video.

We review established theory on modal vibrations, and connect this theory 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. 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 (Figure 2).

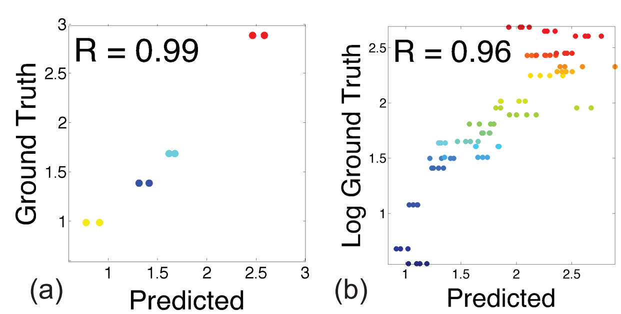

Figure 2: **Example material prediction results**, for the rods (a) and fabrics (b) shown in Figure 1. (a) The rods' Young's moduli as reported by the manufacturer are plotted against values estimated using our technique. (b) Comparisons between ground truth area weight of fabrics and our predictions from videos of those fabrics excited by sound. Each circle in the plots represent the estimated properties from a single video. Colors correspond to Figure 1.

* Joint first author.
This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.