Efficient Perception of Physical Object Properties with Visual Heuristics

Vivian C. Paulun[1,2,3*], Florian S. Bayer[3], Joshua B. Tenenbaum[1], and Roland W. Fleming[3,4]

[1]Department of Brain and Cognitive Sciences, MIT, [2]McGovern Institute for Brain Research, MIT, [3]Department of Experimental Psychology, Justus Liebig University Giessen, [4]Center for Mind, Brain and Behavior (CMBB), University of Marburg and Justus Liebig University Giessen

In order to catch, pick up, or stack objects, i.e., to interact with the physical world around us, we need to accurately infer their physical properties by sight. The computational basis of this ability is still poorly understood. In fact, robust and flexible inferences of physical properties from visual observation are one of the biggest challenges to state-of-the-art artificial vision systems. Here, we propose a model that is based on the observable behavior of objects and materials, i.e., how they typically move or flow. We test this hypothesis for the case of estimating the elasticity of bouncing objects. This is a particularly interesting example because every object can in principle produce an infinite number of trajectories depending on many factors besides elasticity, e.g., initial speed. Yet, in previous work, we have shown that humans accurately estimate the elasticity based on the motion trajectory[1]. Here, we simulated and analyzed the trajectories of 100,000 bouncing cubes and identified 23 motion features the brain could potentially use to estimate elasticity either individually or in combination. In a series of experiments, we carefully teased apart this set of competing but correlated hypotheses. Our results suggest that although humans represent bouncing objects in terms of several different motion features, they rely on single features or heuristics when tasked with estimating elasticity. Interestingly, we found that the stimulus itself determines the heuristic that is used. Specifically, we found that observers rely on the duration of the observed bouncing movement whenever they are able to observe the complete trajectory until the object comes to rest. If, however, the motion was artificially cut short, observers instead used the maximal bounce height to judge elasticity. Thus, our results support the view of a computationally rational observer that uses robust and efficient heuristics when they suffice for the task at hand. Furthermore, we provide evidence for how such heuristics may be derived from observation alone.

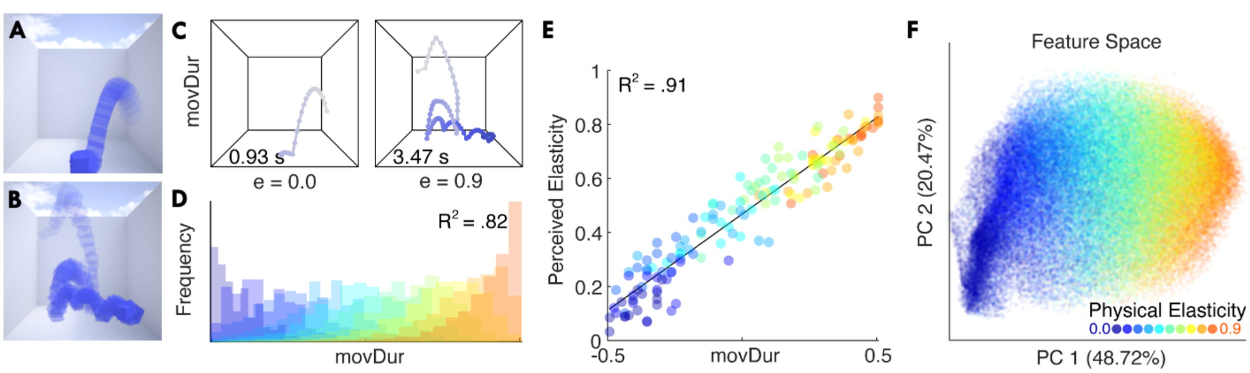

Figure 1. Example stimuli (movie frames overlaid) for low (**A**) and high (**B**) elasticity. **C**) Movement duration of the same two examples and **D**) distribution of that feature over 100k exemplars (R^2 in terms of physical elasticity, which is color-coded). **E**) Movement duration is the best predictor of perceived elasticity, better than physical elasticity which is color-coded. F) 100k stimuli in the space of the first two principal components of all motion features. Stimuli are colored according to their physical elasticity.

---

[1] Paulun VC & Fleming RW (2020). Visually inferring elasticity from the motion trajectory of bouncing cubes. *Journal of Vision, 20(6):6*, 1–14, https://doi.org/10.1167/jov.20.6.6