

# Causal and compositional generative models in online perception

**Ilker Yildirim**

MIT

**Michael Janner**

MIT

**Mario Belledonne**

MIT

**Christian Wallraven**

Korea University

**Winrich Freiwald**

Rockefeller University

**Josh Tenenbaum**

MIT

**Abstract:** From a quick glance or the touch of an object, our brains map sensory signals to scenes composed of rich and detailed shapes and surfaces. Unlike the standard approaches to perception, we argue that this mapping draws on internal causal and compositional models of the physical world and these internal models underlie the generalization capacity of human perception. Here, we present a generative model of visual and multisensory perception in which the latent variables encode intrinsic (e.g., shape) and extrinsic (e.g., occlusion) object properties. Latent variables are inputs to causal models that output sense-specific signals. We present a recognition network that performs efficient inference in the generative model, computing at a speed similar to online perception. We show that our model, but not alternatives, can account for human performance in an occluded face matching task and in a visual-to-haptic face matching task.