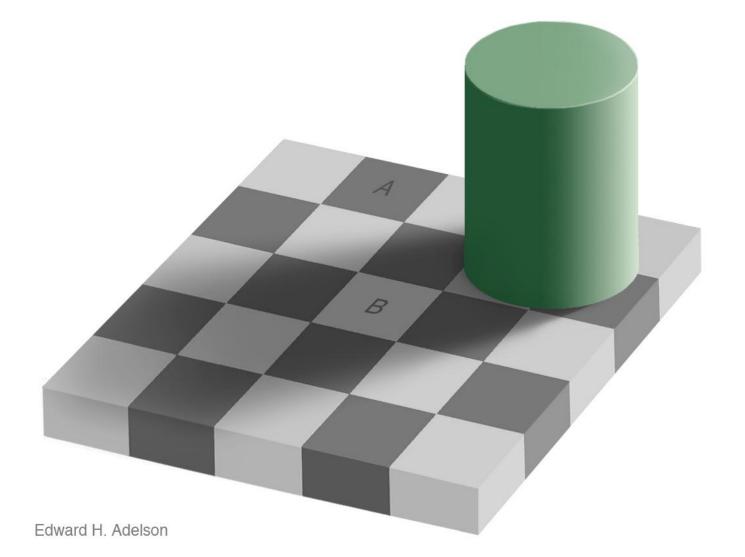
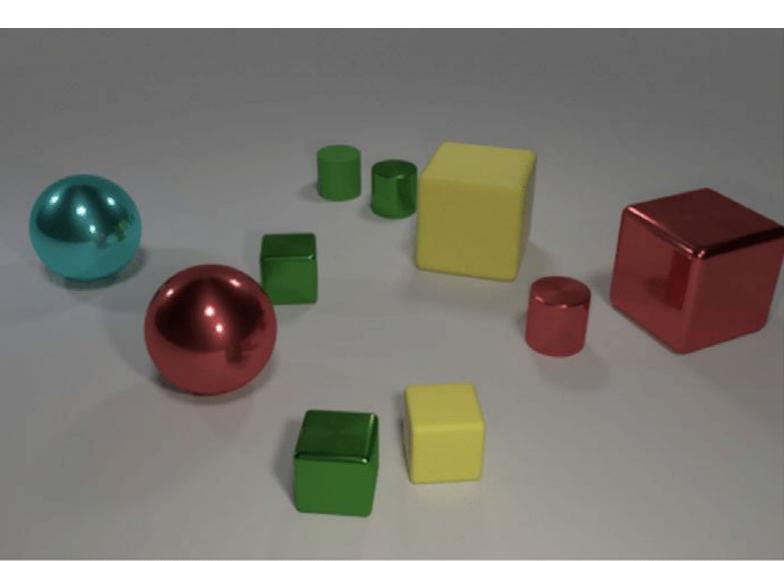
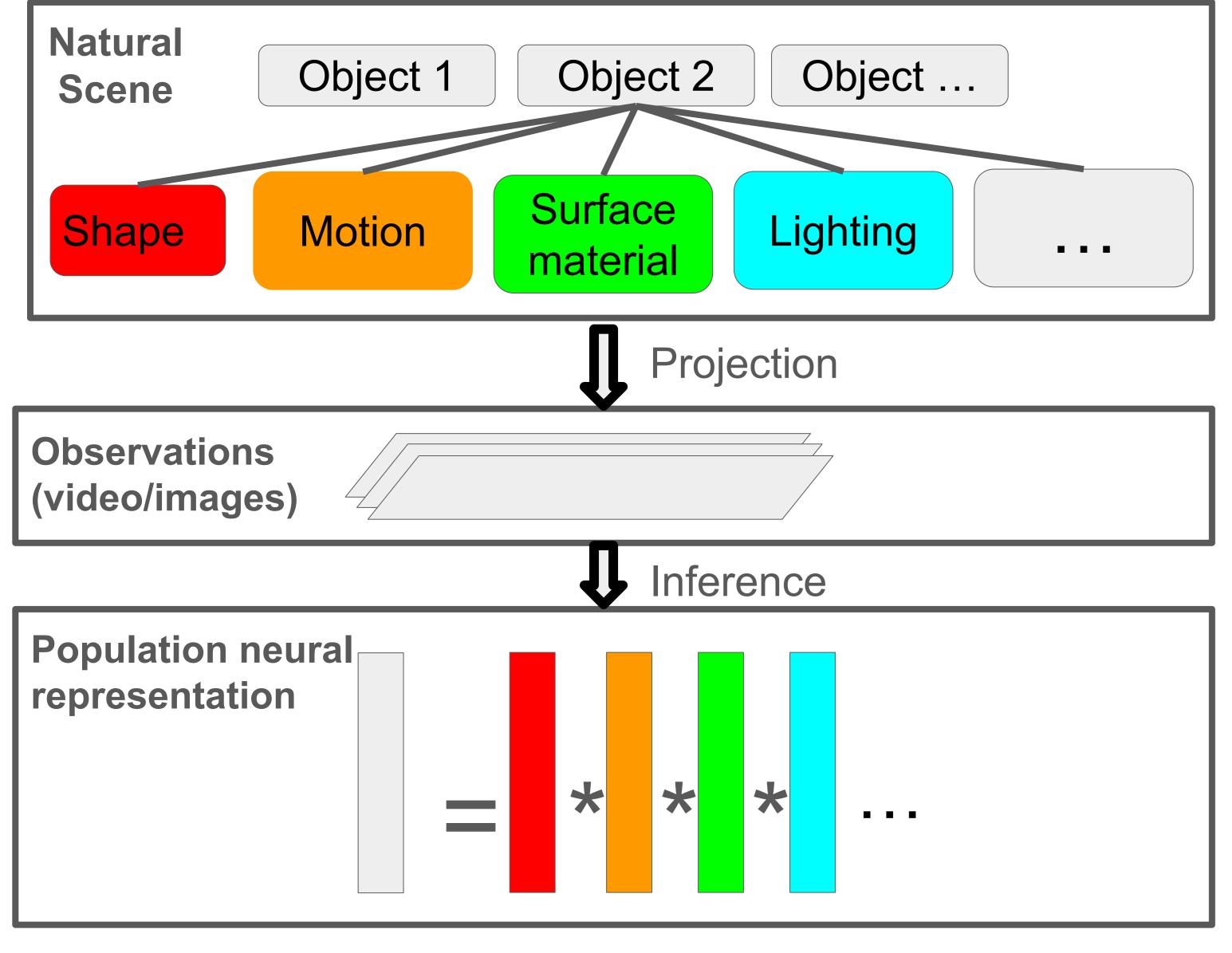
#### General goal of visual cortex

General goal: learns to form a representation of these external causes from raw sensory input. Specifically, scene are composed with objects with natural different factors. E.g. shape, reflectance, position, objects motion, . . .





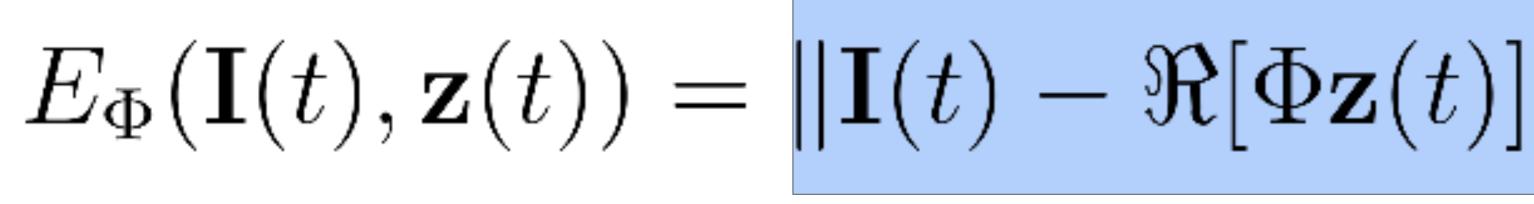
#### **Efficient factorizable representation**



In this work, we focus on modeling the video composed of objects and transformation, by factorizing form and motion.

## **Energy function**

s.t.



### preserving information sparseness

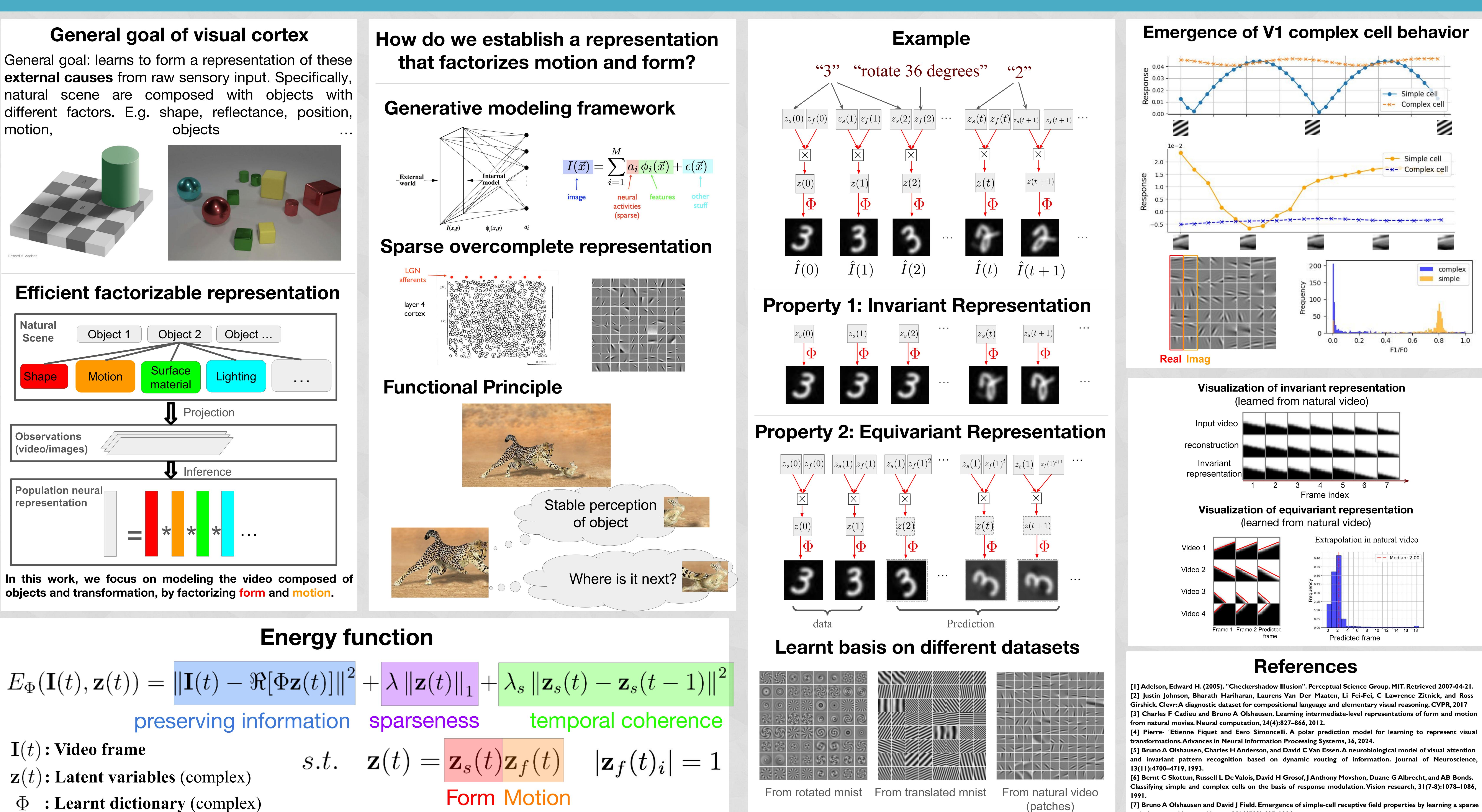
I(t): Video frame

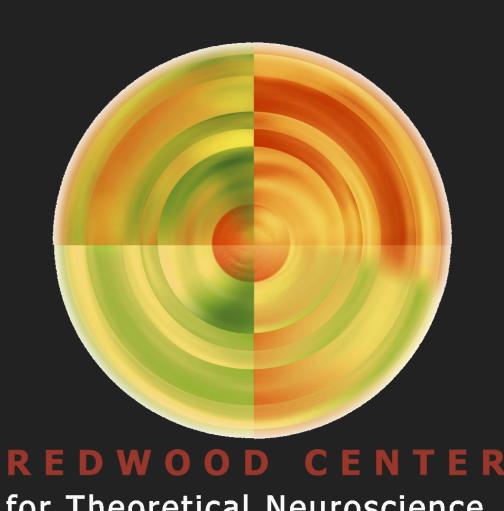
z(t): Latent variables (complex)

 $\Phi$ : Learnt dictionary (complex)

# Predictive and Invariant Representations via Motion and Form Factorization in Natural Scenes Zeyu Yun<sup>1,3,5</sup>, Christopher Kymn<sup>3,4</sup>, Galen Chuang<sup>2,3</sup>, Yubei Chen<sup>5</sup>, Bruno Olshausen<sup>1,2,3,4</sup>

(1) EECS Dept. (2) Vision Science (3) Redwood Center for Theoretical Neuroscience (4) Helen Wills Neuroscience Inst., UC Berkeley (5) ECE Dept. UC Davis





for Theoretical Neuroscience

[7] Bruno A Olshausen and David J Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature, 381(6583):607, 1996.