Recognizing and Learning Object Categories: Year 2009

Li Fei-Fei, Stanford

Rob Fergus, NYU

Antonio Torralba, MIT

Testimonials: “since I attended this course, I can recognize all the objects that I see”

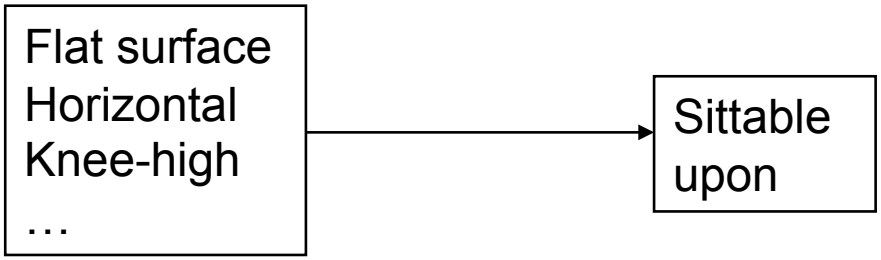
Why do we care about recognition?

Perception of function: We can perceive the 3D shape, texture, material properties, without knowing about objects. But, the concept of category encapsulates also information about what can we do with those objects.

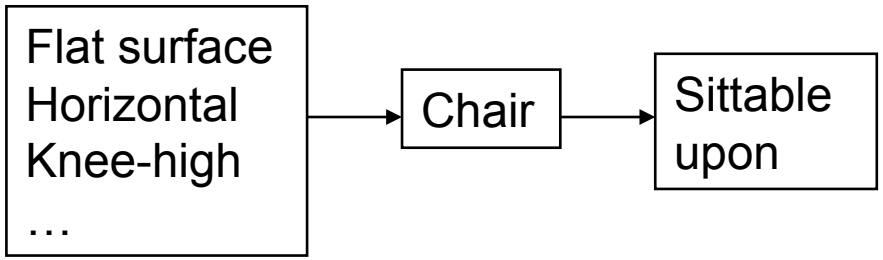
"We therefore include the perception of function as a proper –indeed, crucial- subject for vision science", *from Vision Science, chapter 9, Palmer*.

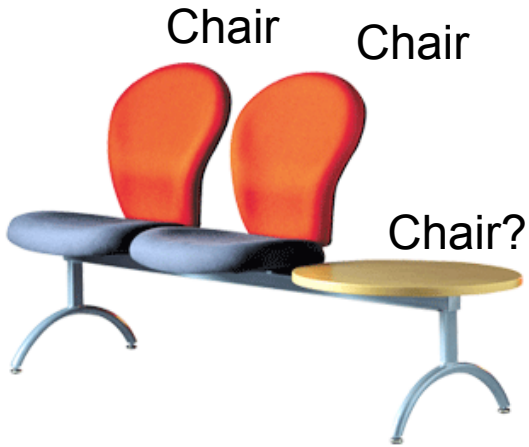
The perception of function

- Direct perception (affordances): Gibson



- Mediated perception (Categorization)

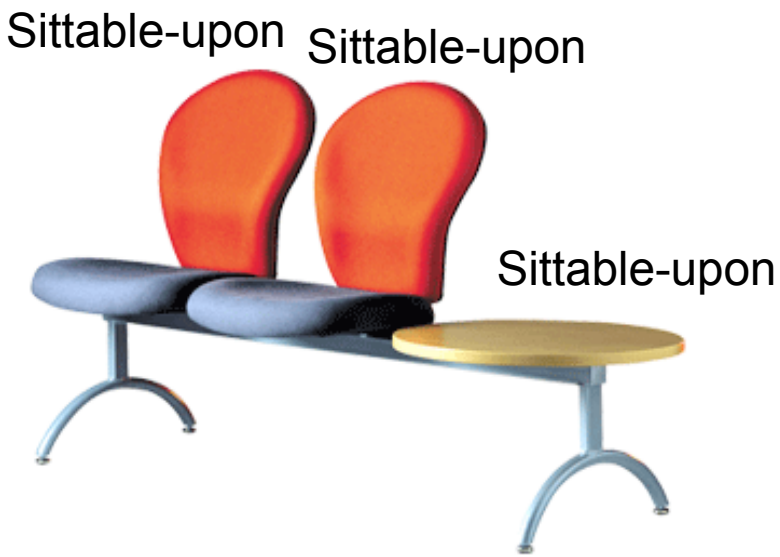




Direct perception

Some aspects of an object function can be perceived directly

- Functional form: Some forms clearly indicate to a function (“sittable-upon”, container, cutting device, …)



It does not seem easy to sit-upon this…

Direct perception

Some aspects of an object function can be perceived directly

- Observer relativity: Function is observer dependent

Limitations of Direct Perception

Objects of similar structure might have very different functions

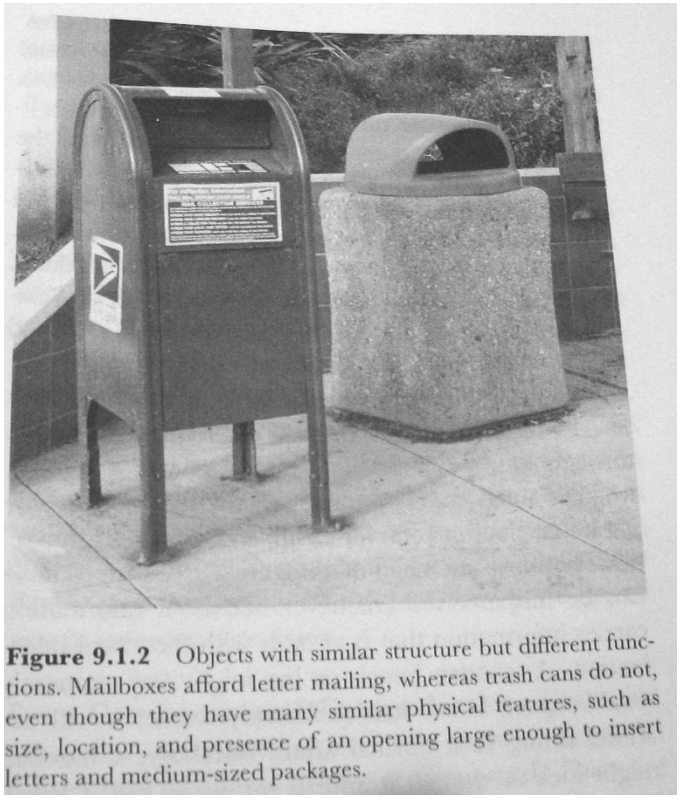


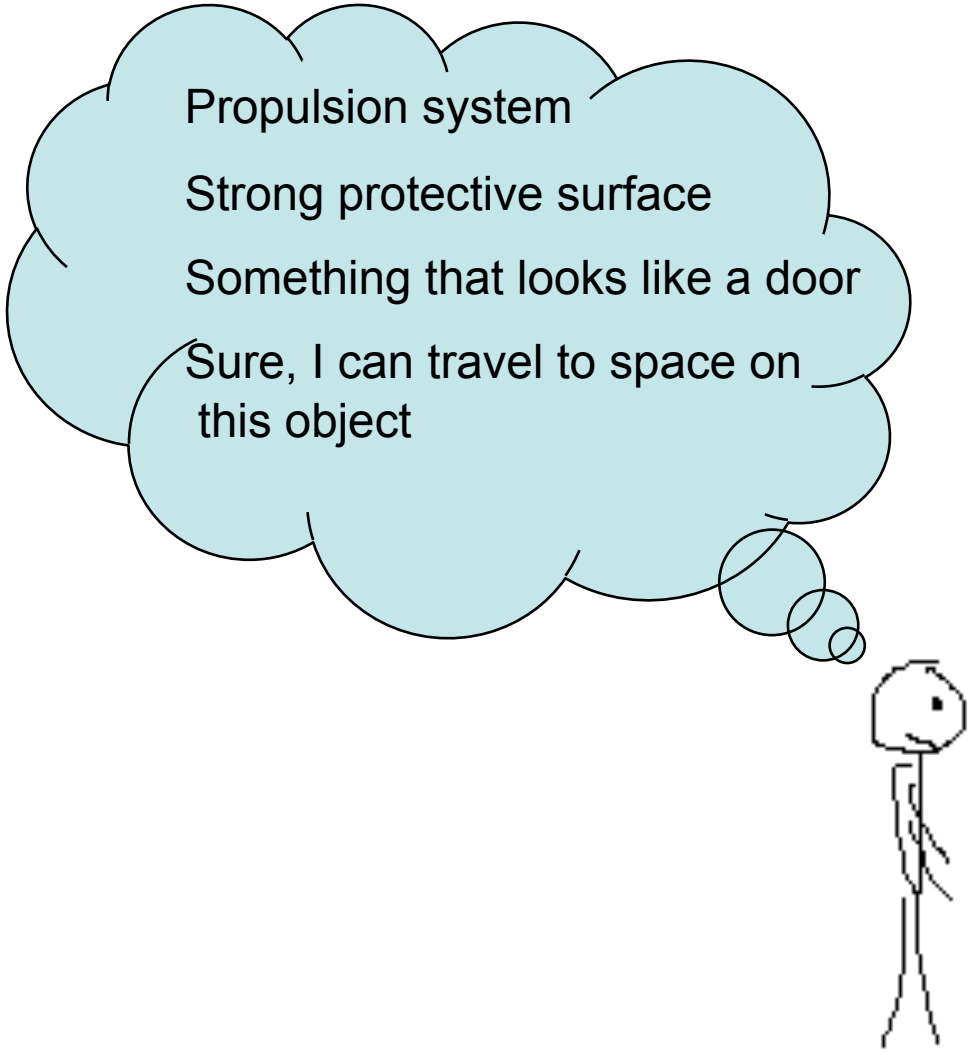
Figure 9.1.2 Objects with similar structure but different functions. Mailboxes afford letter mailing, whereas trash cans do not, even though they have many similar physical features, such as size, location, and presence of an opening large enough to insert letters and medium-sized packages.

Not all functions seem to be available from direct visual information only.

The functions are the same at some level of description: we can put things inside in both and somebody will come later to empty them. However, we are not expected to put inside the same kinds of things…

Limitations of Direct Perception

Visual appearance might be a very weak cue to function

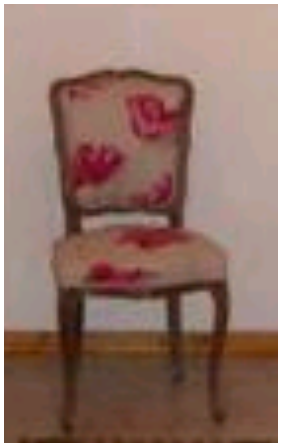


How do we achieve Mediated perception?

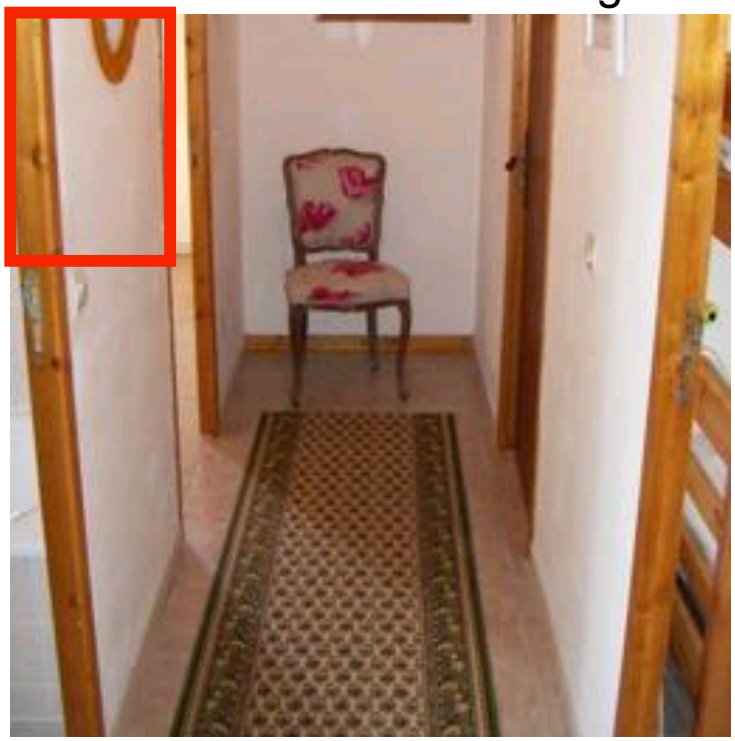
Well… this requires object recognition (for more details, see entire course)

Object recognition
Is it really so hard?

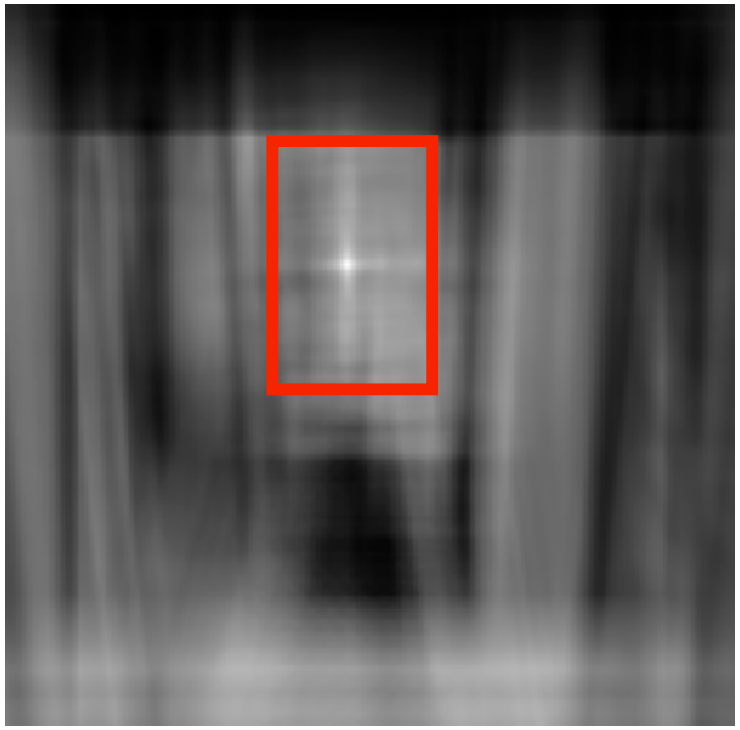
This is a chair

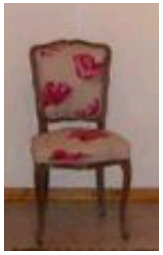


Find the chair in this image



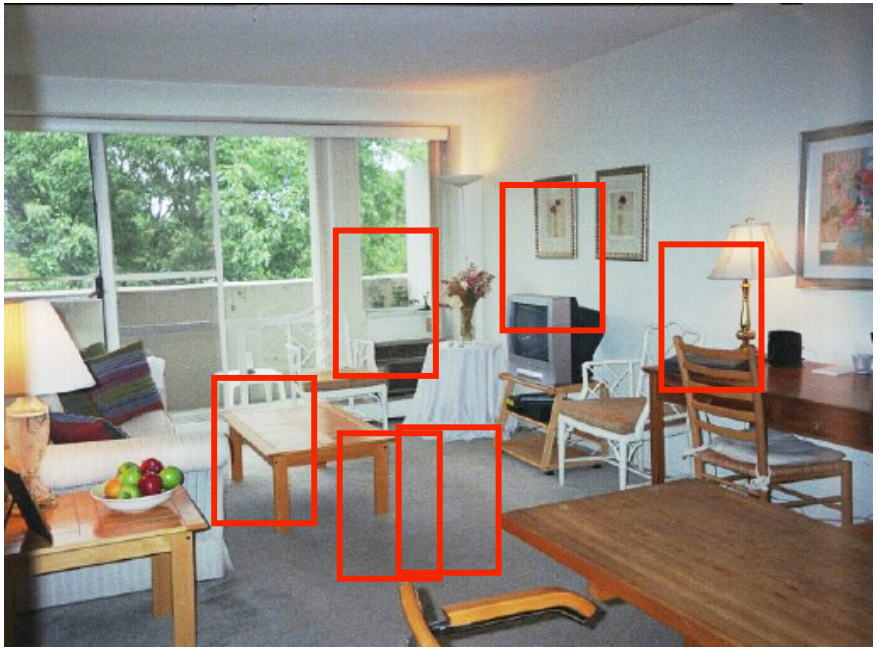
Output of normalized correlation

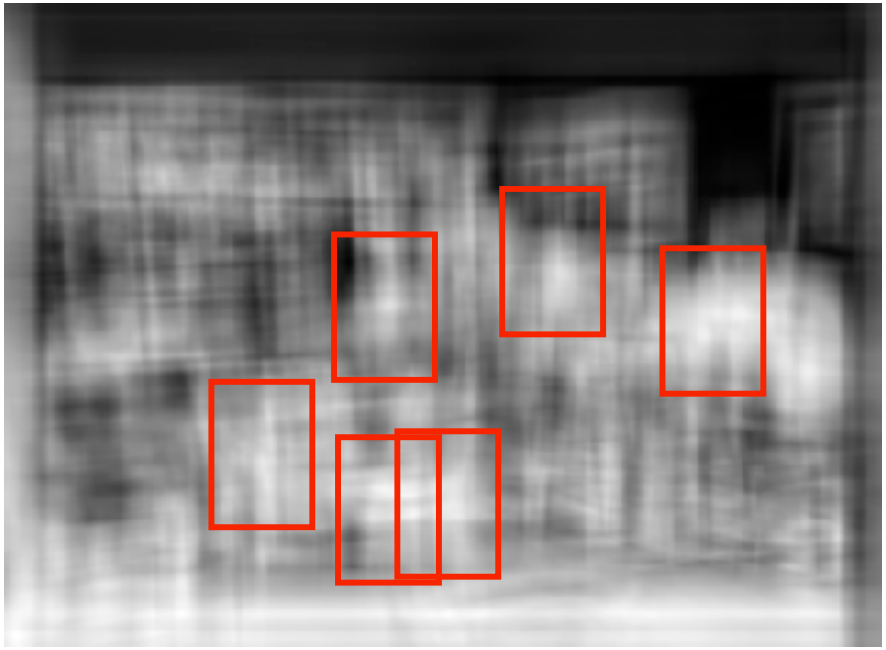




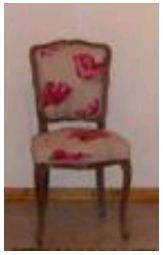
Object recognition
Is it really so hard?

Find the chair in this image





Pretty much garbage
Simple template matching is not going to make it



Object recognition
Is it really so hard?

Find the chair in this image

A “popular method is that of template matching, by point to point correlation of a model pattern with the image pattern. These techniques are inadequate for three-dimensional scene analysis for many reasons, such as occlusion, changes in viewing angle, and articulation of parts.” Nivatia & Binford, 1977.

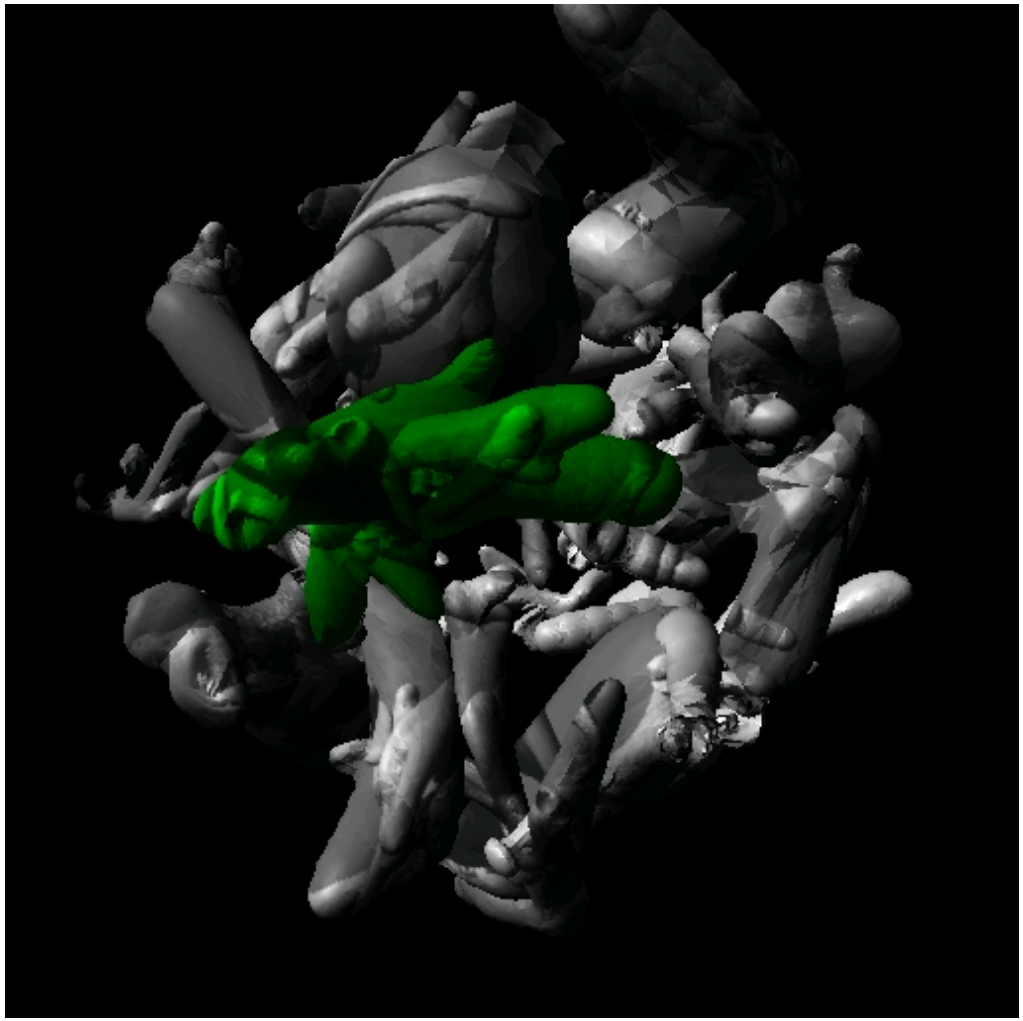
And it can get a lot harder

Brady, M. J., & Kersten, D. (2003). Bootstrapped learning of novel objects. J Vis, 3(6), 413-422

your visual system is amazing

is your visual system amazing?

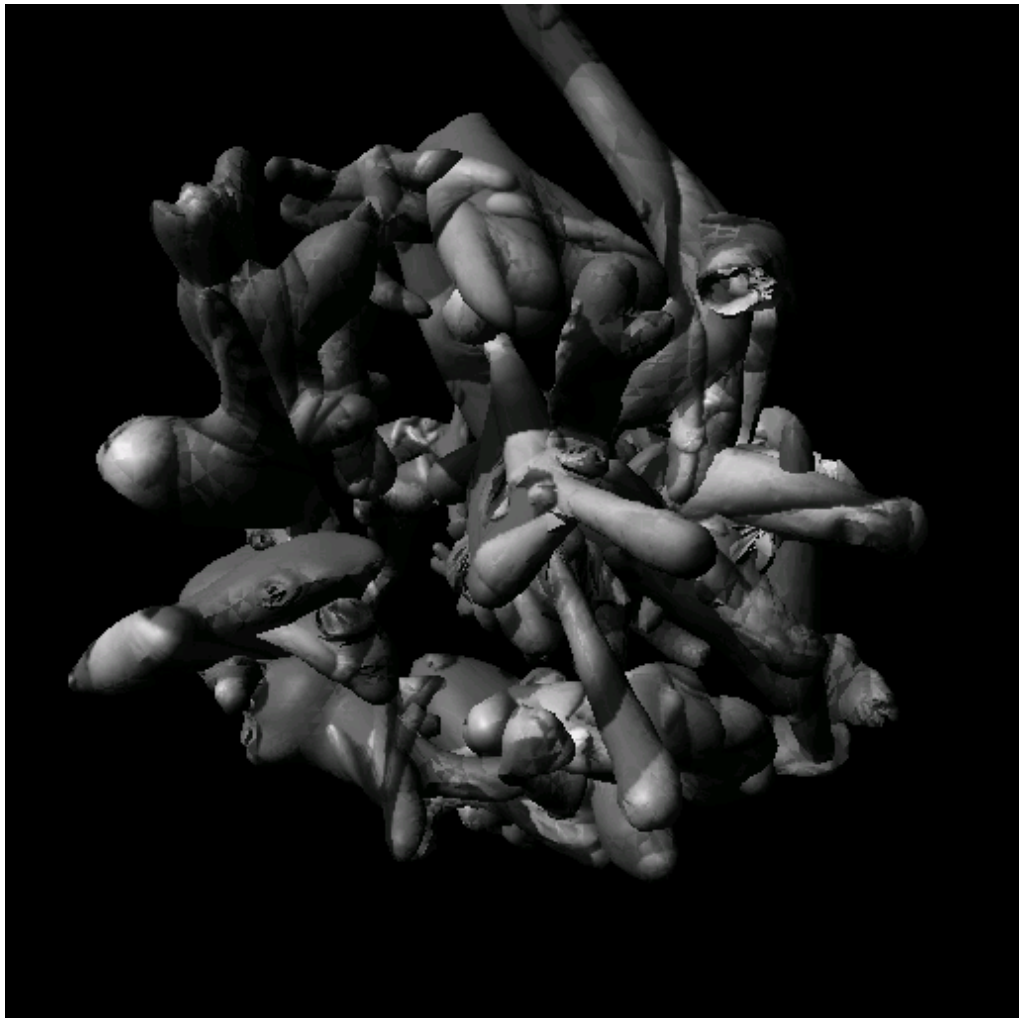
Discover the camouflaged object



Brady, M. J., & Kersten, D. (2003). Bootstrapped learning of novel objects. J Vis, 3(6), 413-422

Discover the camouflaged object

Brady, M. J., & Kersten, D. (2003). Bootstrapped learning of novel objects. J Vis, 3(6), 413-422



Any guesses?

Outline

1. Introduction (15')

2. Single object categories (1h15')

    - Bag of words (rob)
    - Part-based (rob)
    - Discriminative (rob)
    - Detecting single objects in contexts (antonio)
    - 3D object classes (fei-fei)

15:30 – 16:00 Coffee break

3. Multiple object categories (1h30')

    - Recognizing a large number of objects (rob)
    - Recognizing multiple objects in an image (antonio)
    - Objects and annotations (fei-fei)

4. Object-related datasets and challenges (30')