CSC2503: Foundations of Computer Vision

Object Recognition

Most slides are modified from the excellent course notes and tutorials by Antonio Torralba, Fei-Fei Li and Rob Fergus.
http://people.csail.mit.edu/torralba/cvpr2007/

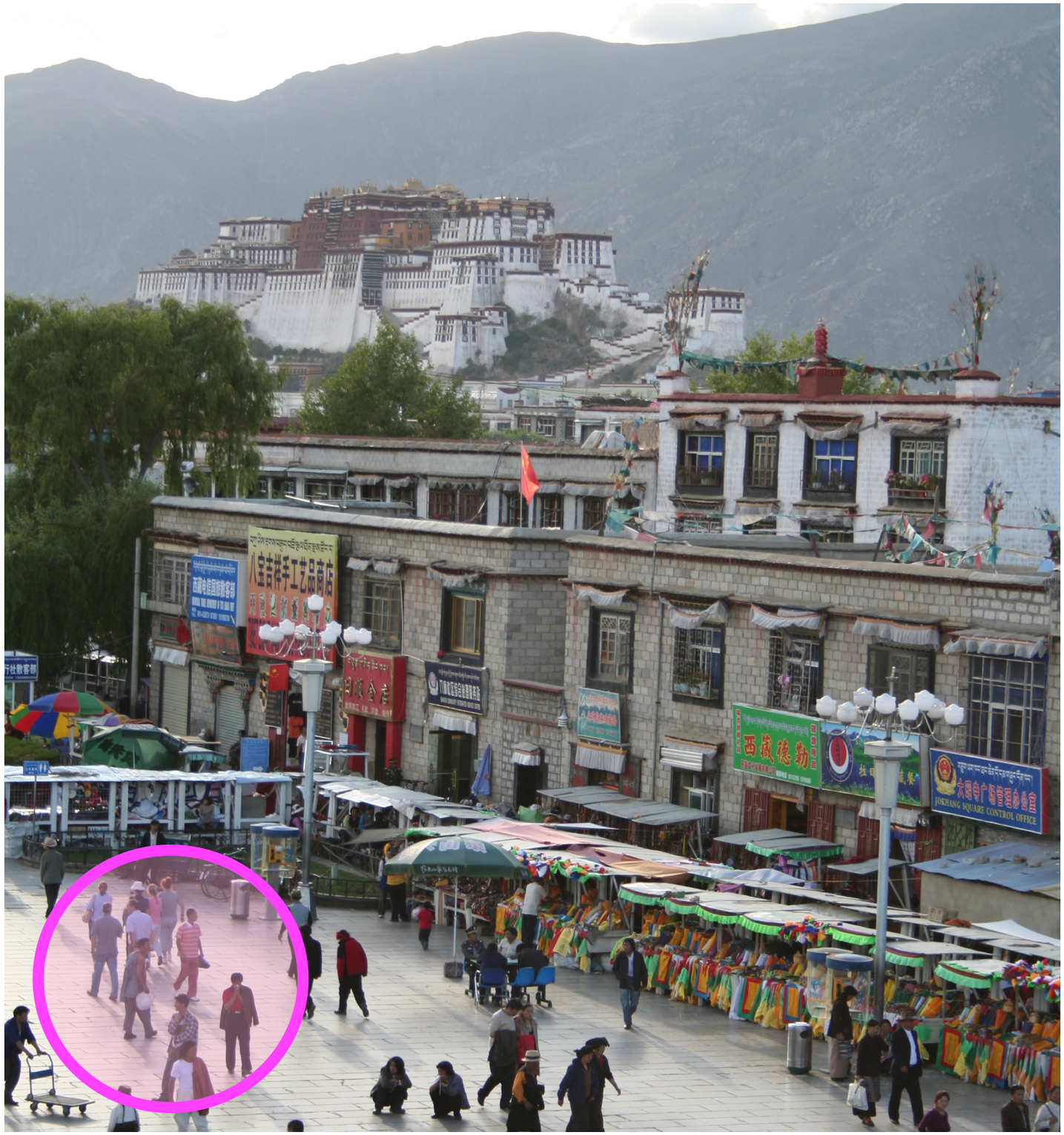
What’s involved in visual recognition?

Verification

Is this a lamp?

Detection

Are there people in the image?



Identification

Is this Potala Palace?

Category recognition

Scene and context categorization

- outdoor
- city
- daytime
- ...

Viewpoint and space

Are the distances large or small?

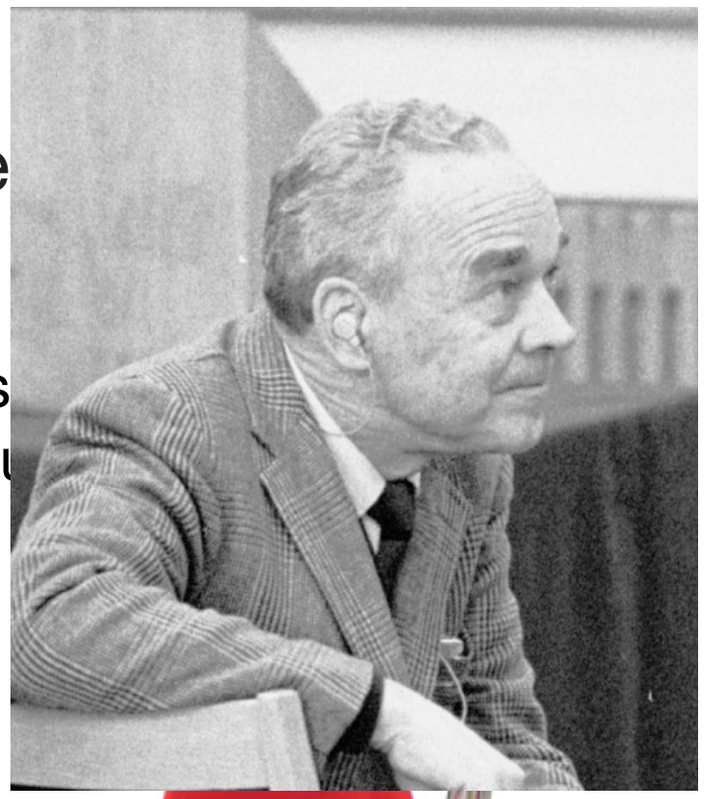
How far are the distant buildings?

Activity recognition

Perception, categories and function?

- Gestalt perception (1920-40)
- Direct perception & affordances (Gibson 1950-60)
- Mediated perception (categorization)

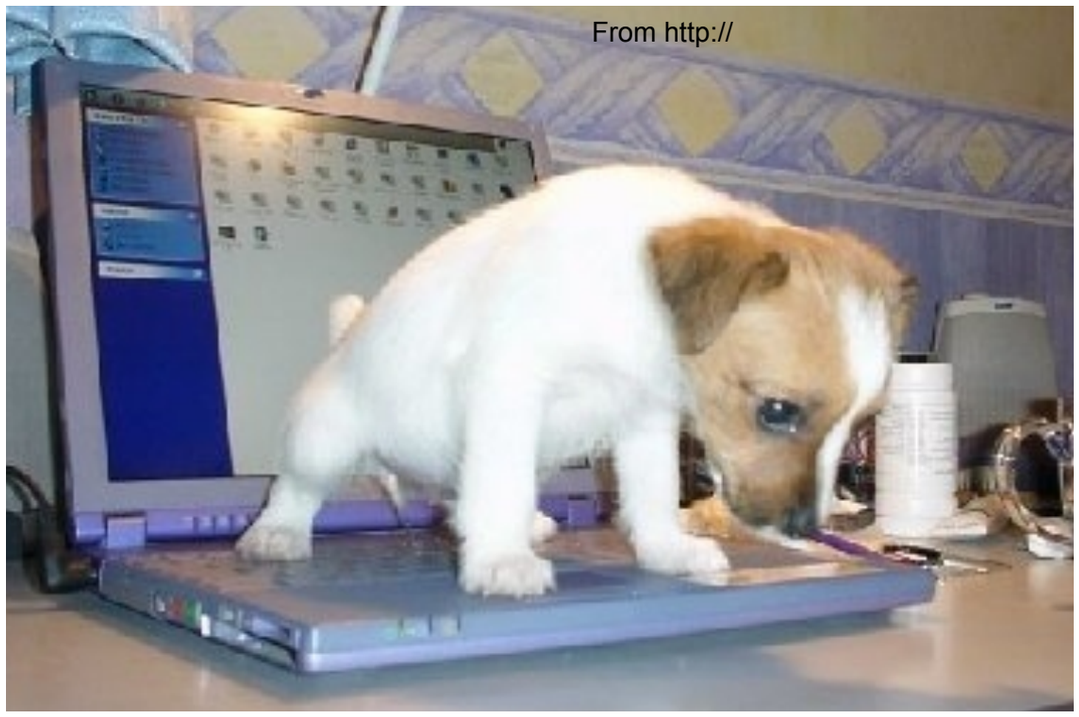
Some aspects of object function can be
(container, supporting surface, ...)



It does
to sit-u

JJ Gibson

Perception, categories and function?



Some aspects of function are observer dependent

Perception, categories and function?

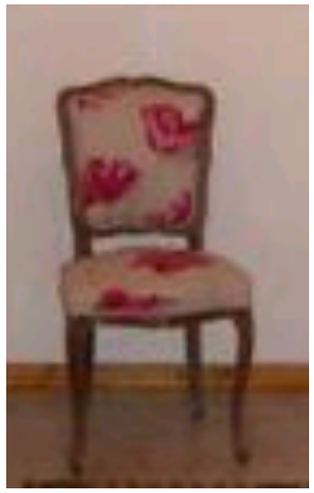
Objects with similar structure might have very different functions

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. Here the functions are the same at some level: 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…

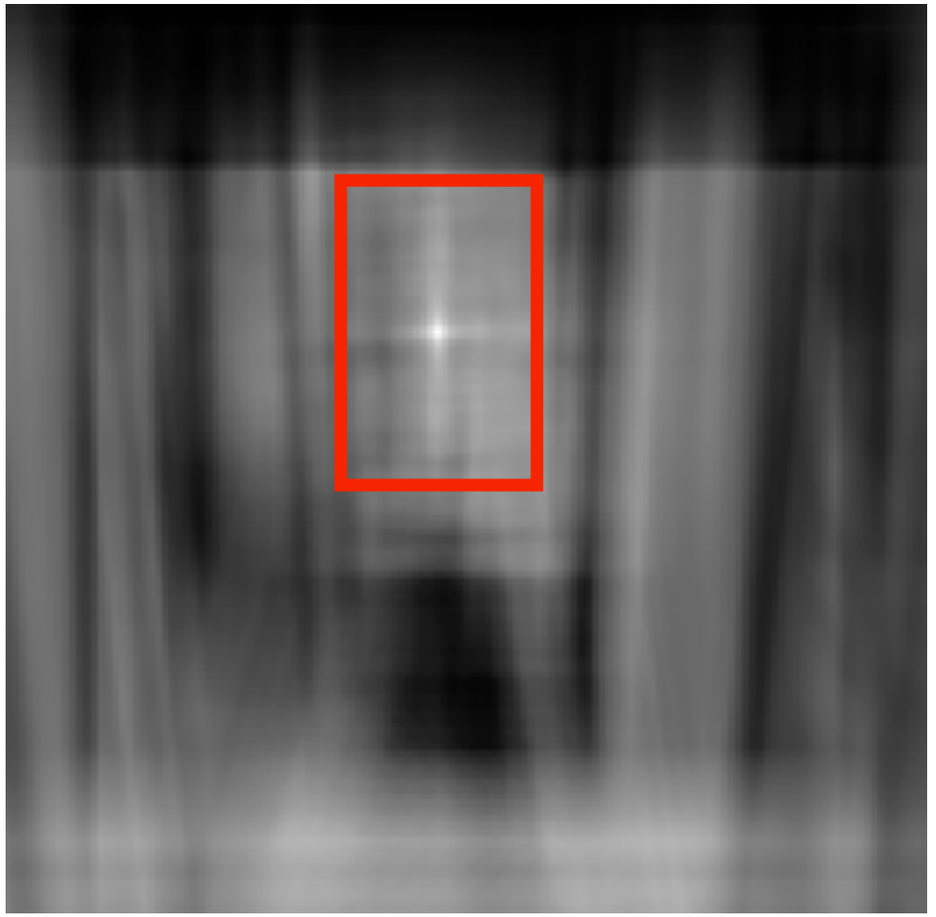
Is recognition really that hard?

This is a chair

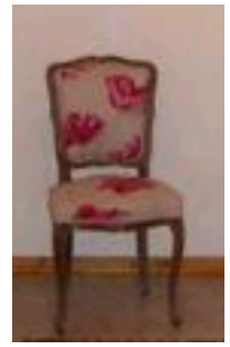


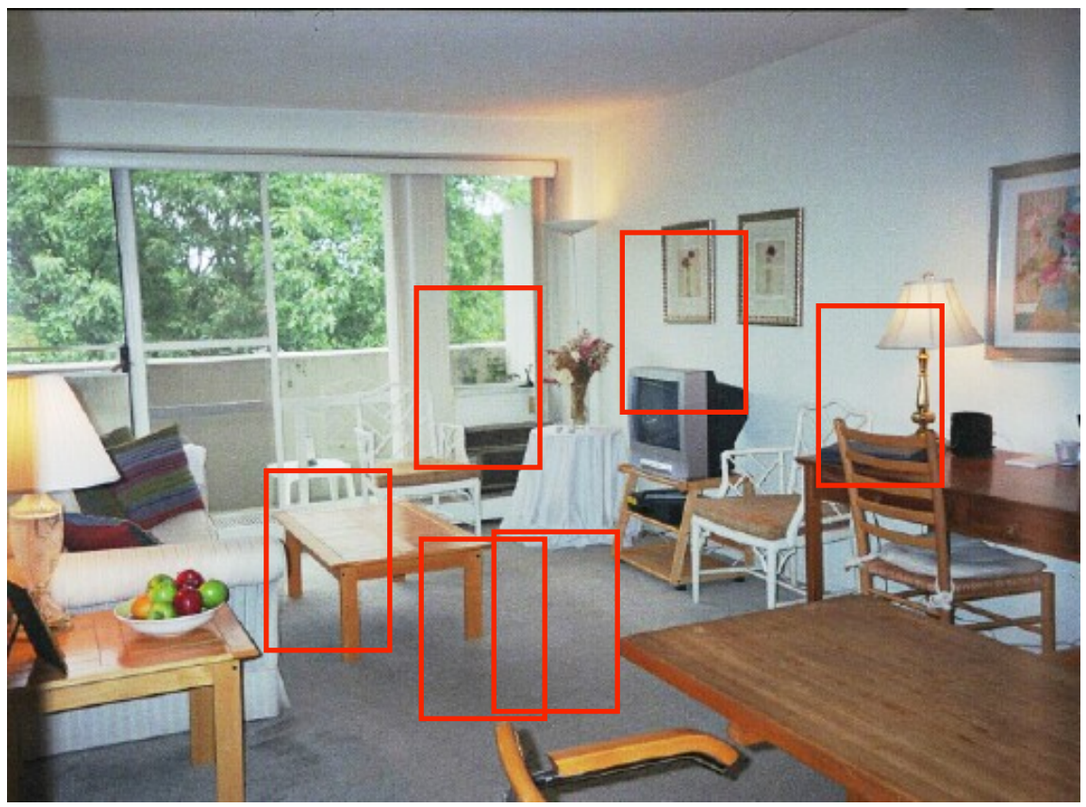
Find the chair in this image

Normalized correlation

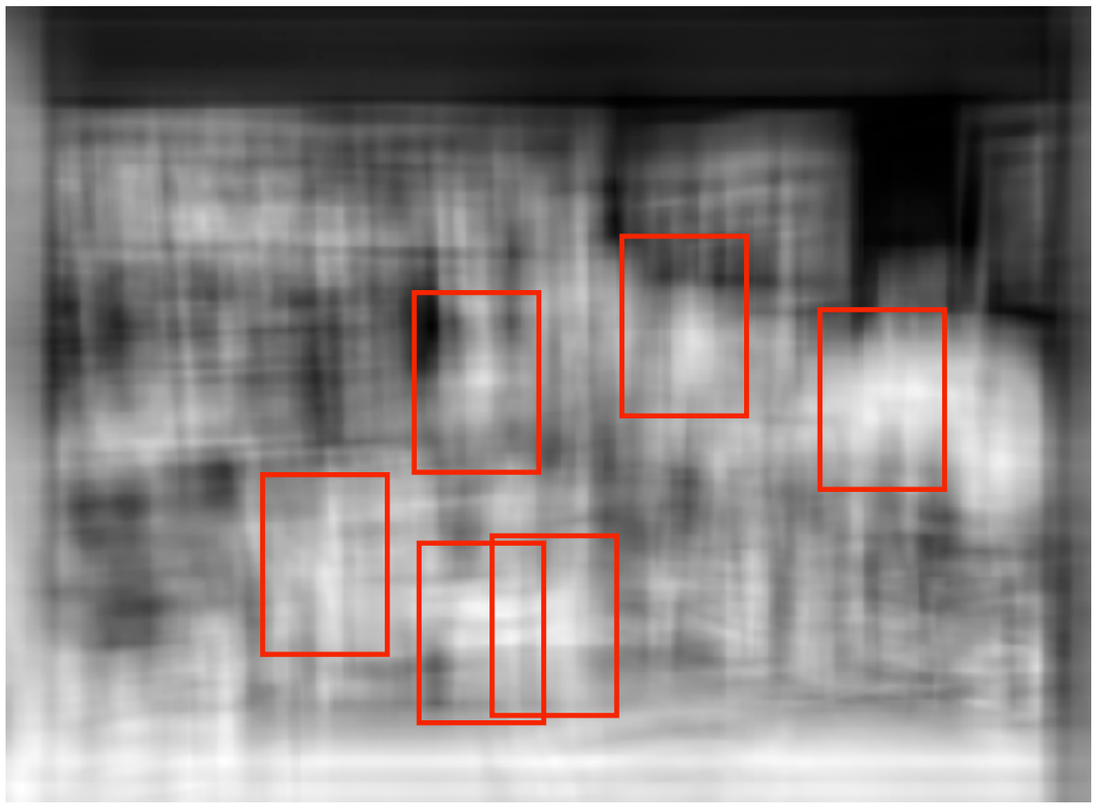


Is recognition really that hard?





Find the chair in this image



Will template matching work?

What makes object recognition hard?

Challenges: Viewpoint

Michelangelo 1475-1564

Challenges: Illumination

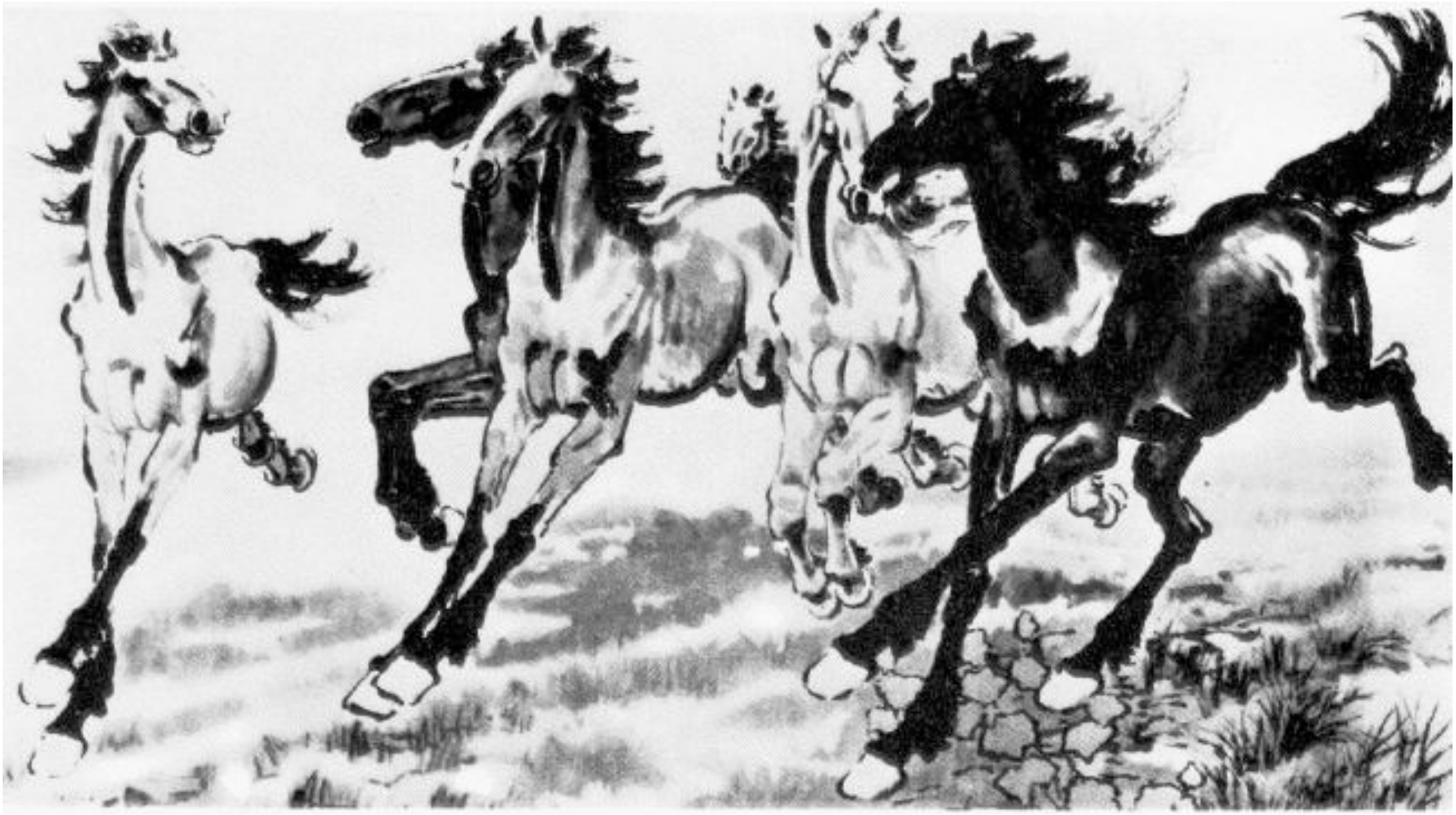
slide credit: S. Ullman

Challenges: Occlusion

Magritte, 1957

Challenges: Scale

Challenges: Deformation



Xu, Beihong 1943

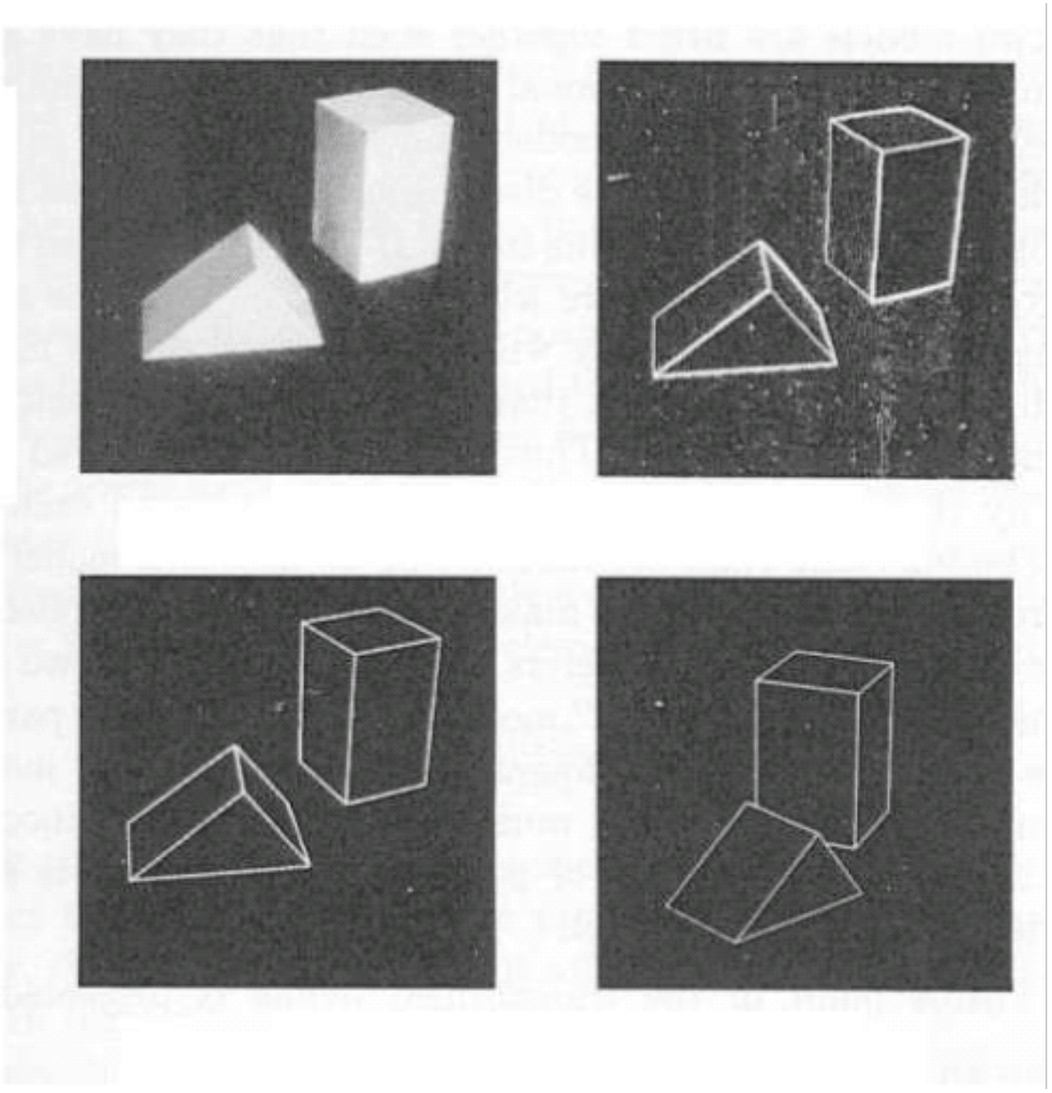
Challenges: Background clutter

Klimt, 1913

Challenges: Intra-class variation

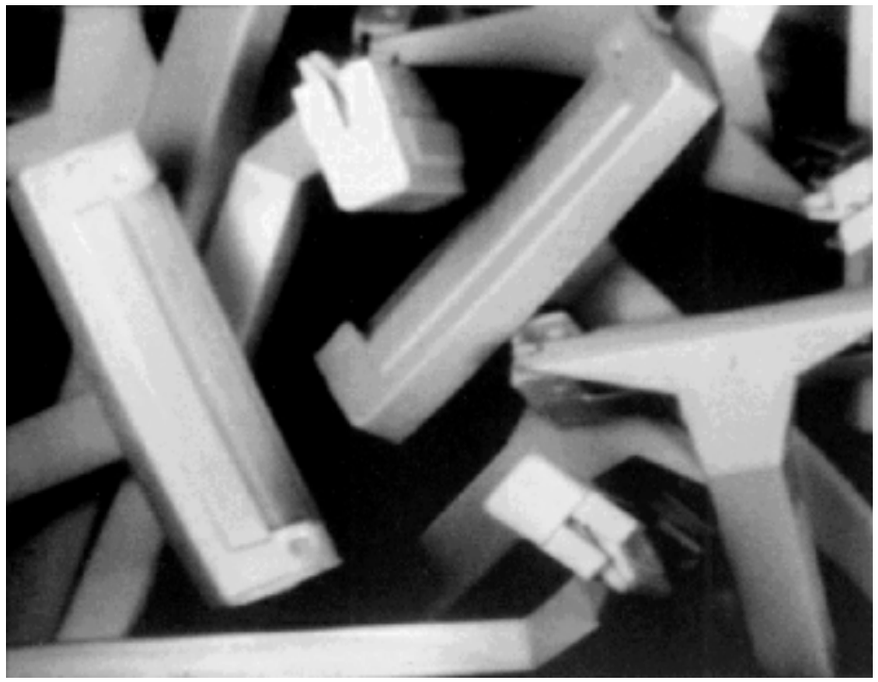
Blocks World

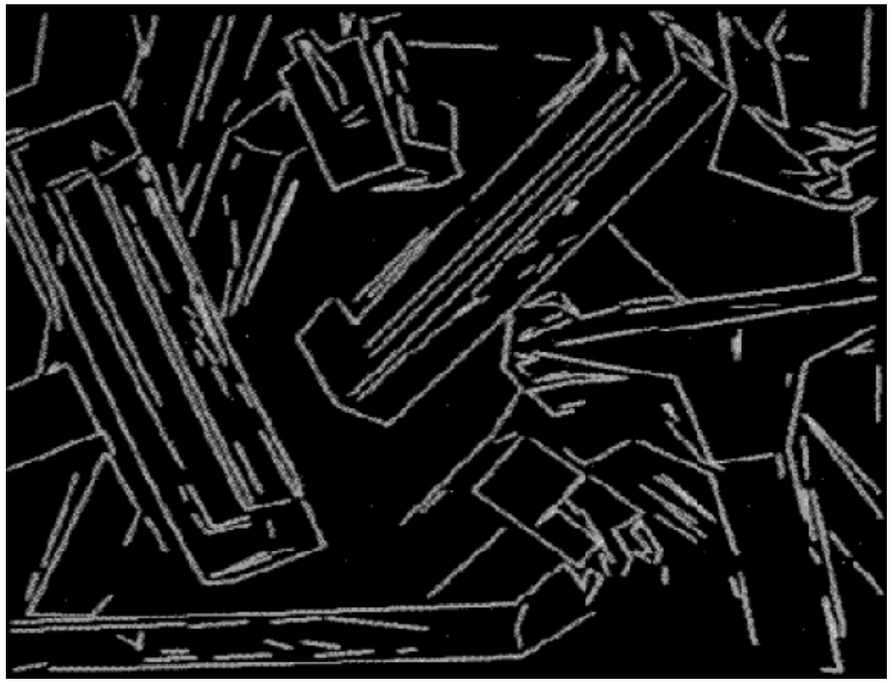
L. Roberts

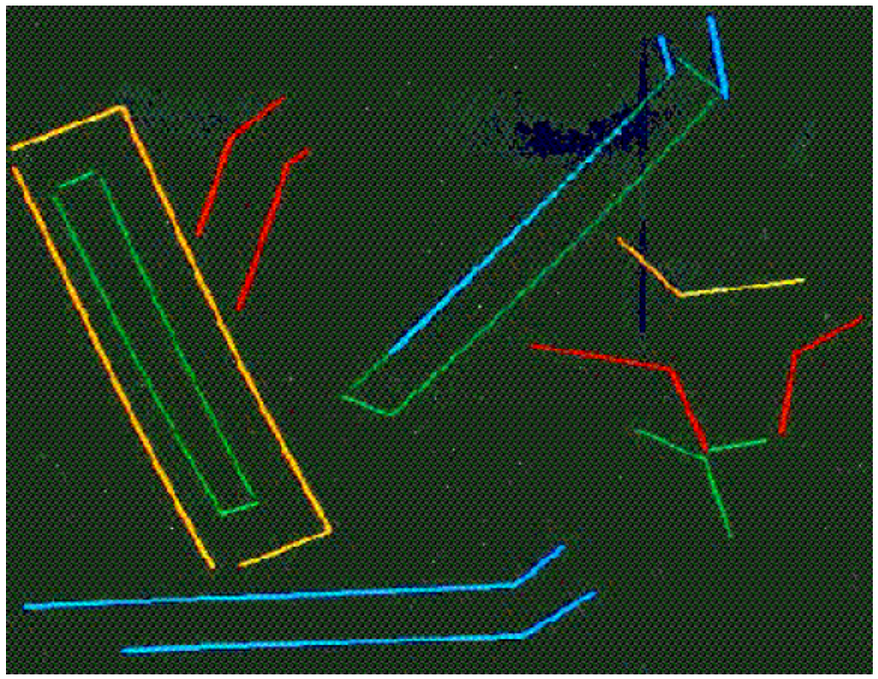


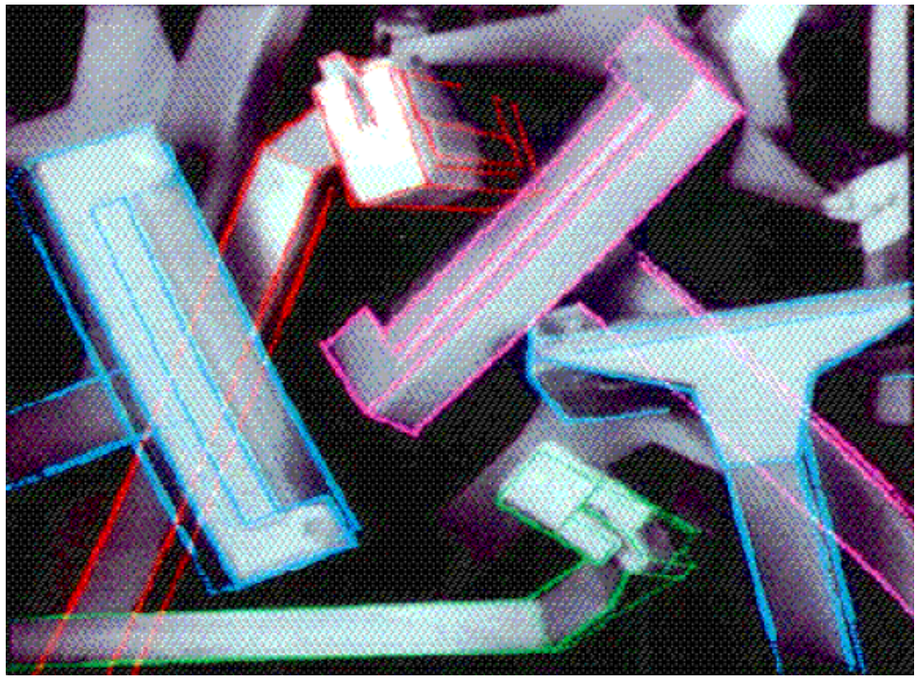
1960s-70s: Constrained 3D scene models to allow object recognition from very simple image features (Lambertian, trihedral objects), edge labeling (junction analysis), and object recognition.

Geometric Matching





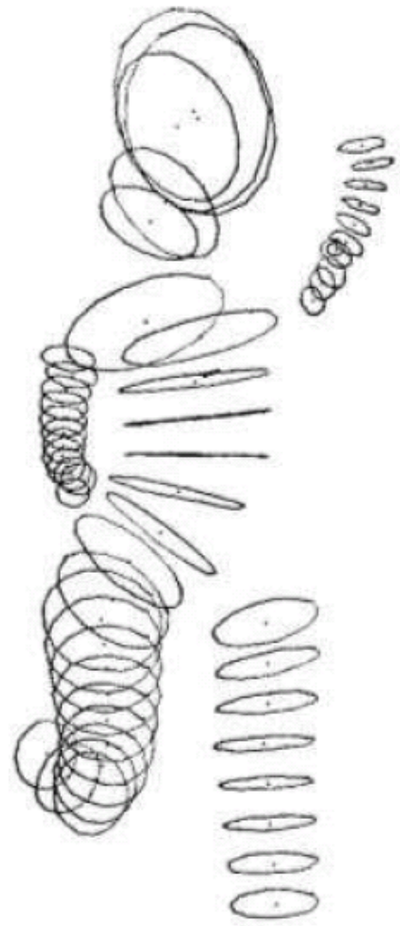




D Lowe circa 1985

Generalized Cylinders

T. Binford

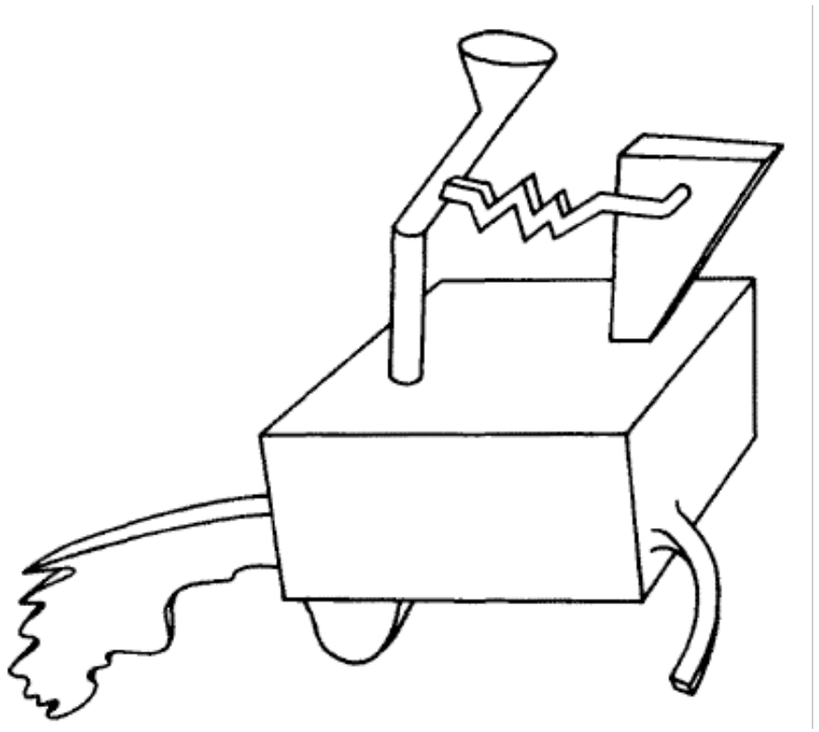


1970s-80s: Designing languages for representing 3D shapes and parts, e.g., in terms of “Generalized Cylinders” (cylinders modulated w/ sweeping rule)

Recognition by parts

I. Biederman

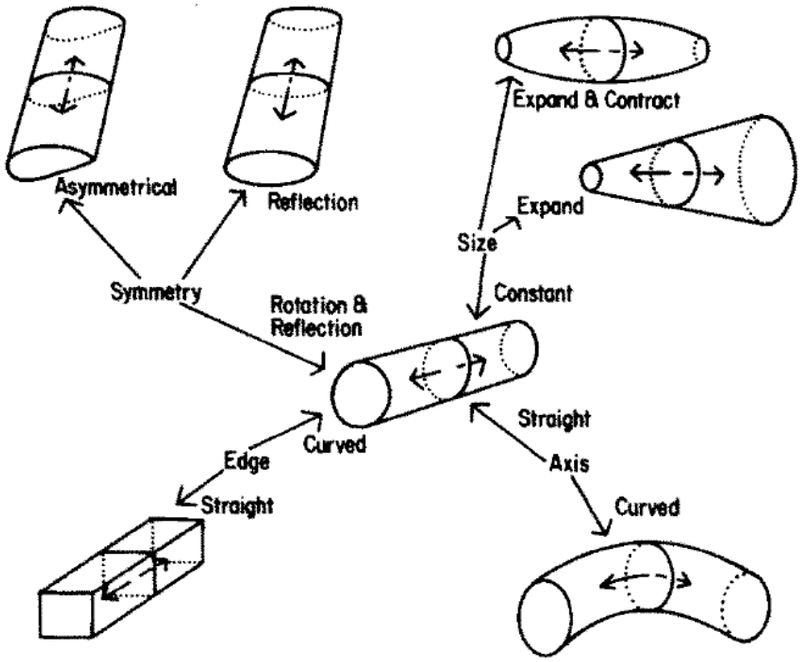
Late 1980s: Vocabulary for shape parts, estimated from images via rules of “perceptual organization” (e.g., collinearity, symmetry, parallelism, ...)



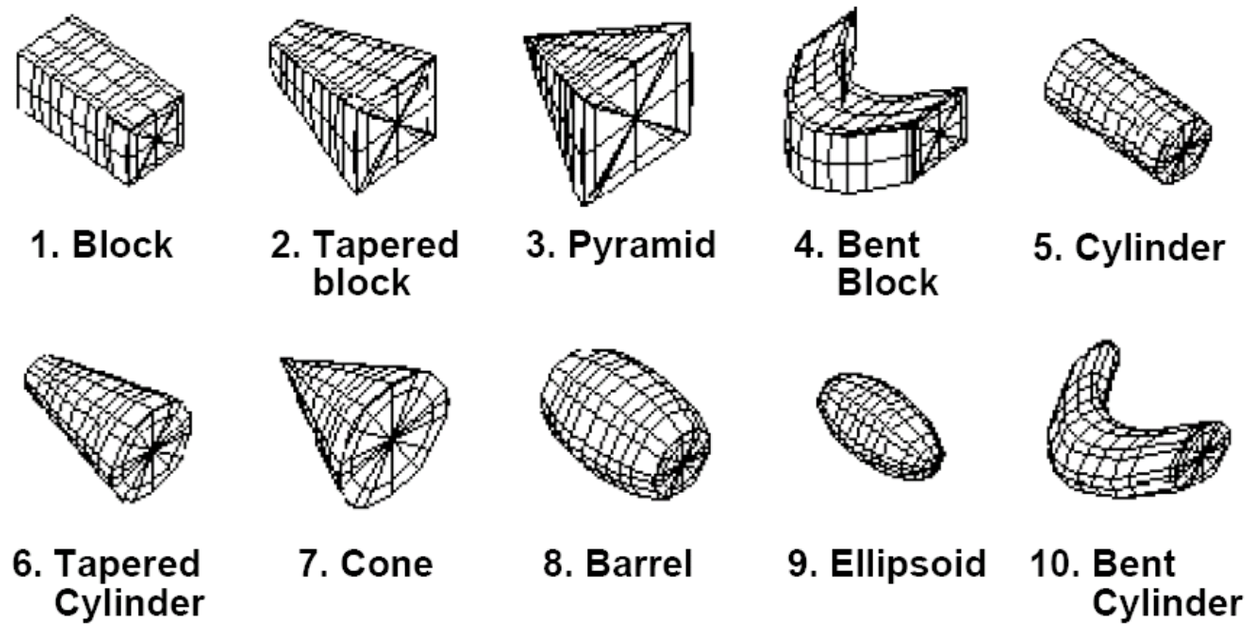
Unfamiliar fictional objects are consistently perceived in terms of parts, with similarity to familiar objects (“a hot dog cart”).

Recognition by parts

Geons: Shape primitives + deformations, with predictable edge properties under perspective.



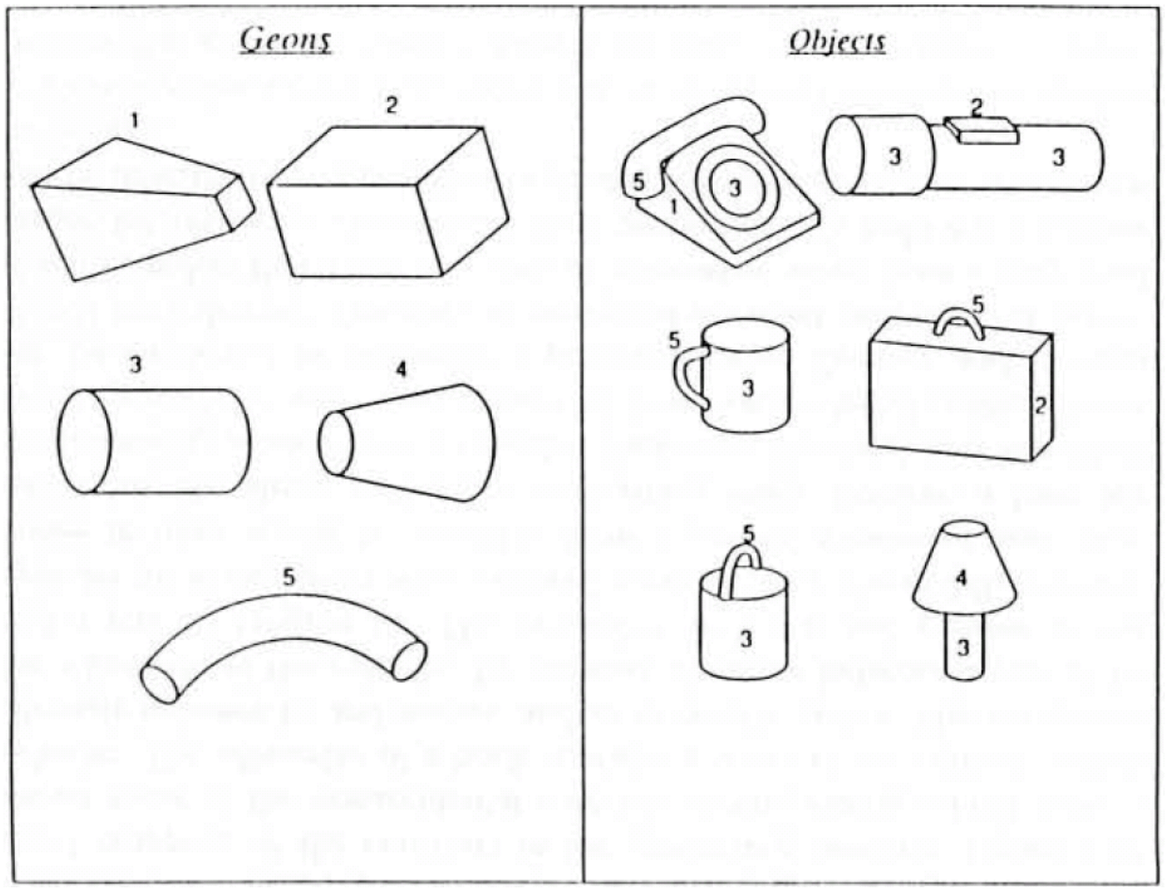
I. Biederman, 1987



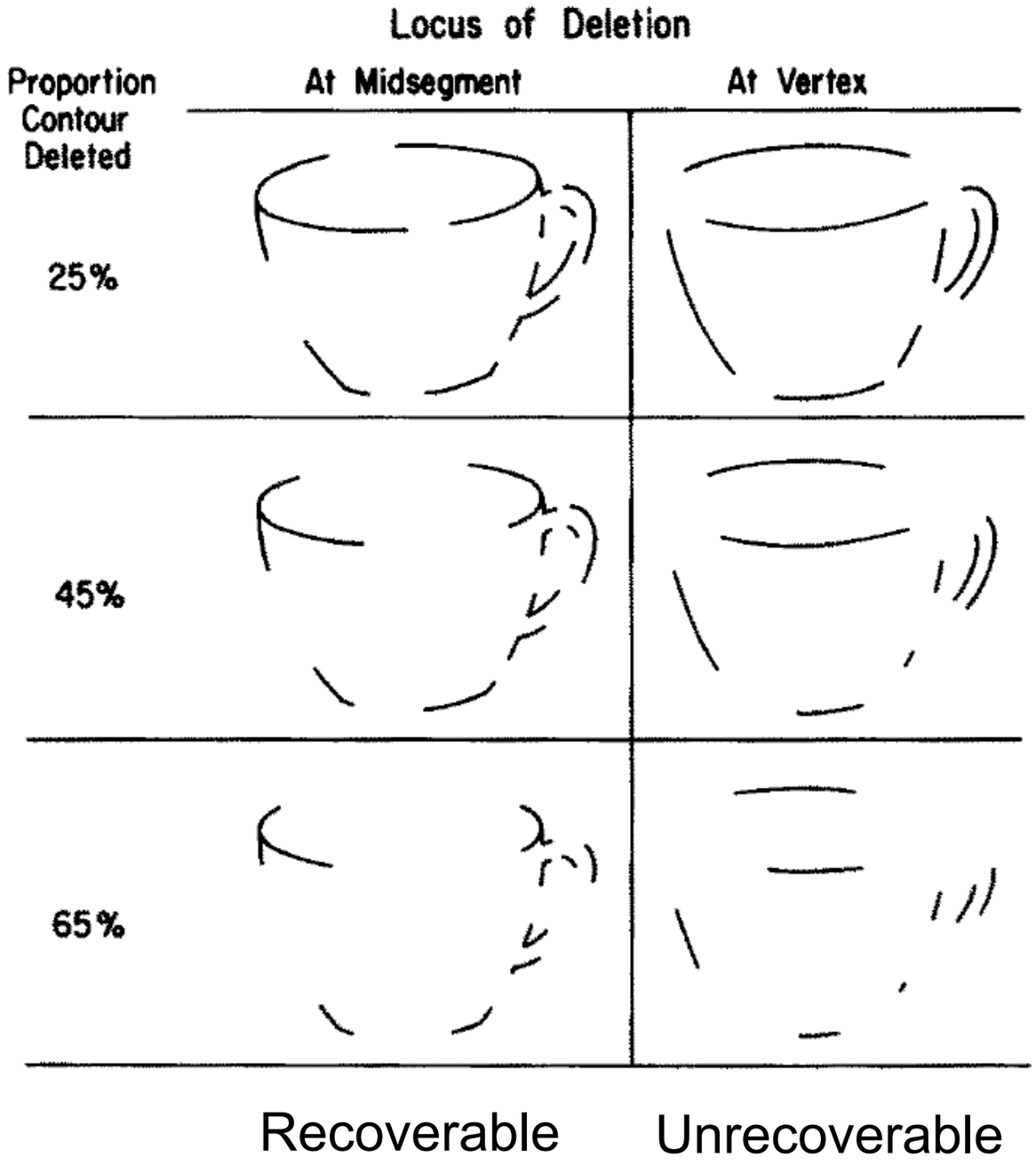
Introduced in vision by A. Pentland, 1986.

Recognition by parts

Geons: Shape primitives + deformations, with predictable edge properties under perspective.



Recognition by parts

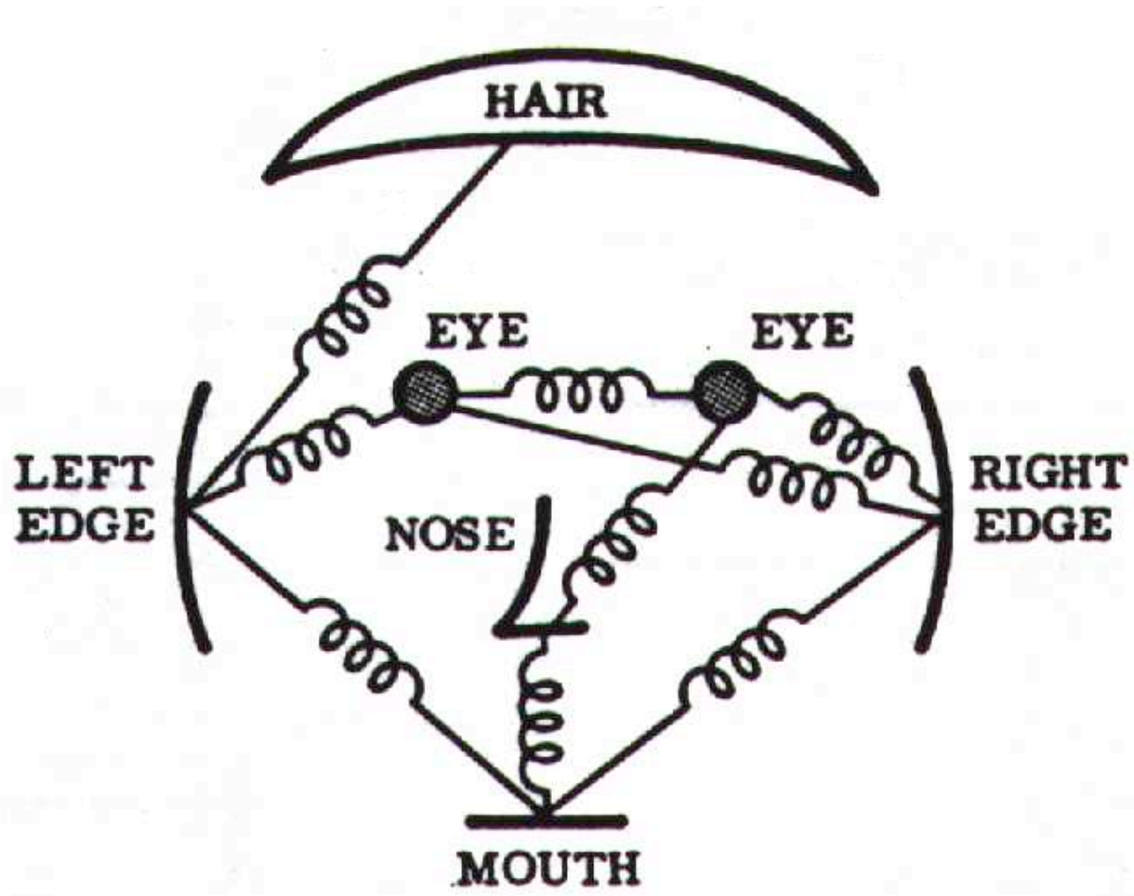


Parsing based on contours, non-accidental properties & concavities

Parts + Spatial Configurations

There is more to shape than just the right part primitives.

Spatial relationships are also important.



[Fischler & Elschlager 73]

View-based recognition

1995+: Lose the 3D ... just find things in 2D.

View-based models

The “Margaret Thatcher Illusion”, by Peter Thompson

View-based models

Turk and Pentland: Face detection using eigen-faces