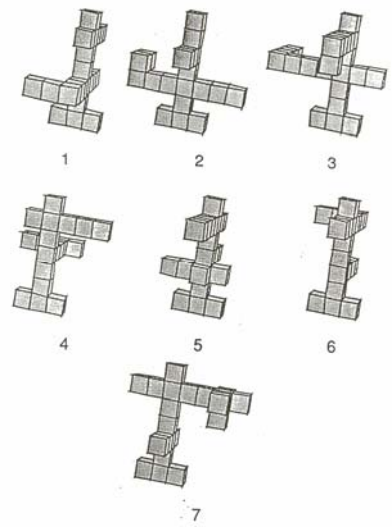
High-Level Vision

Object Recognition III

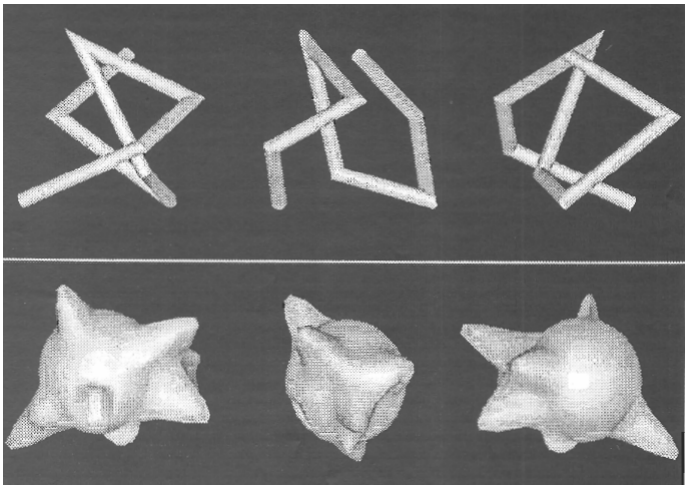
CS332 Visual Processing
Department of Computer Science
Wellesley College

Viewer-centered object representation?

Tarr, '95: After learning to recognize a set of 3-D objects from a small set of specific 2-D views of these objects, the time needed to recognize a novel view is proportional to the 3-D angle between the new view and closest learned view



Bulthoff & Edelman



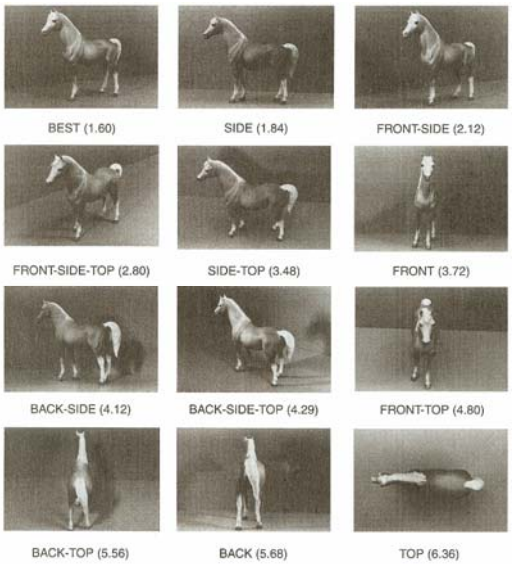
1-3

Canonical views

Palmer et al.:

- subjective ratings
- response times to name object category

Strong correlation in performance of two tasks



1-4

Canonical views

Best views of common objects, based on subjective ratings

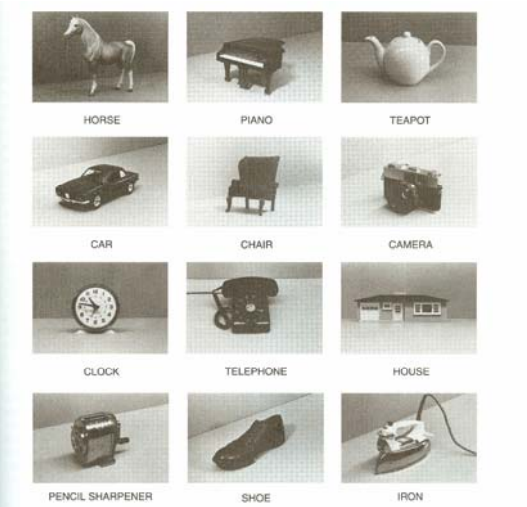


Figure 9.2.5 Canonical perspective for 12 objects. The perspective view receiving the best rating is shown for each of the 12 objects in the study by Palmer, Rosch and Chase (1981). In subsequent studies, subjects named these views more quickly than views from other perspectives, as indicated in Figure 9.2.6. (From Palmer, Rosch, & Chase, 1981.)

1-5

Be careful about viewpoint dependence…

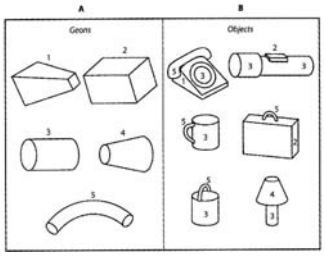
1-6

Parts decomposition

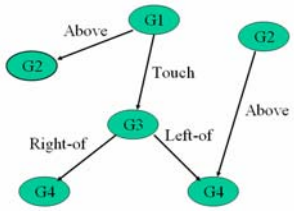
Every instance of each object class shares certain parts, arranged in a certain way

(1) find object parts

(2) recognize objects by presence of parts with proper relationship



Biederman's Geons



"Structural Description"

1-7

Biederman & Gerhardstein

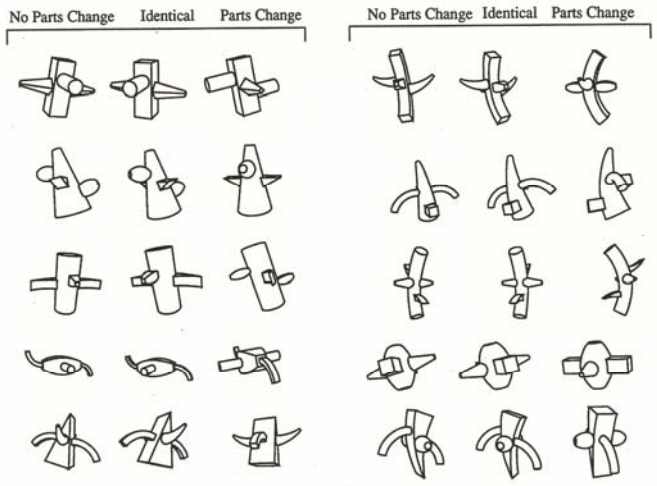
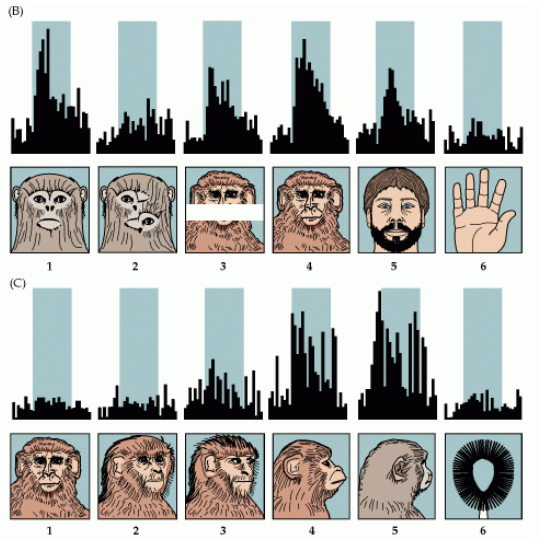
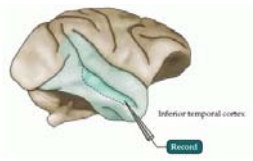


Figure 10. The 10 unfamiliar objects used in Experiment 3. (The no-parts-change and parts-change views are rotations of 45° in depth in different directions from the zero view. Note that no object contains a part unique to that object, and relations between object parts are the same for all objects.)

1-8

Early recognition physiology



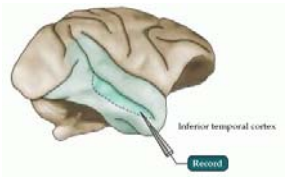


1-9

Logothetis & Pauls

With one learned view, monkeys can generalize to +/- 40° rotation away from this view

With two learned views 75° apart, monkeys recognized all intermediate views (> 95% correct performance)



IT neurons were found that responded selectively to these novel objects (about 10% of ~800 neurons), *most were viewpoint dependent*

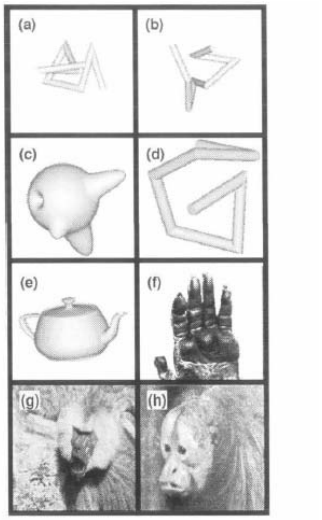
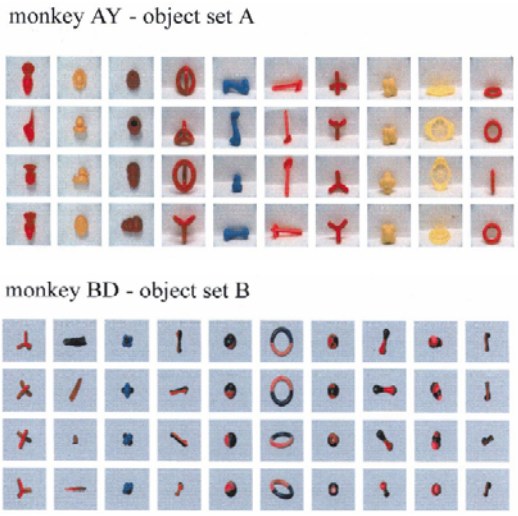


Figure 2. Examples of stimuli used in experiments on object recognition. Wire-like, amoeboid, and common-type objects were created mathematically and rendered by a computer. Pictures of various natural objects such as hands and faces were digitized using a camera and a standard PC-based frame-grabber.

1-10

Booth & Rolls



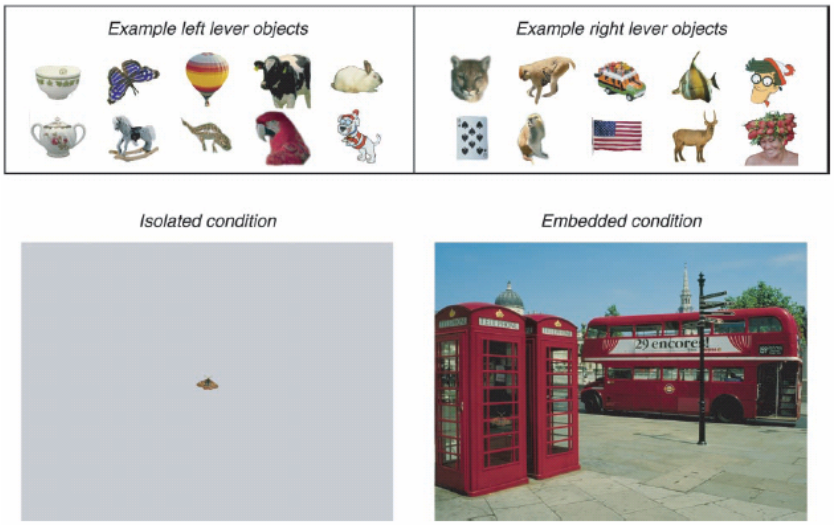
Recorded from IT neurons after monkeys had objects in cages for a few weeks

~ 70 neurons responded to cage objects in *view dependent* way

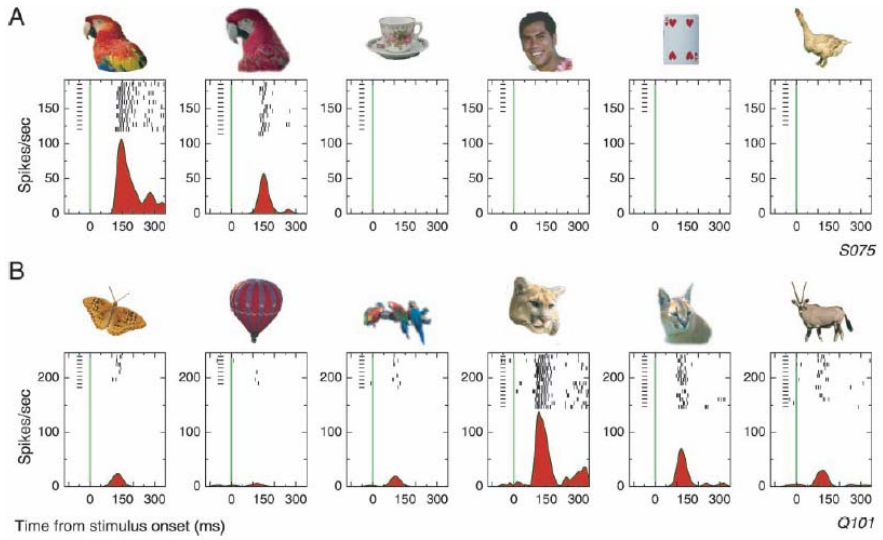
~ 20 neurons responded to cage objects in *view invariant* way

1-11

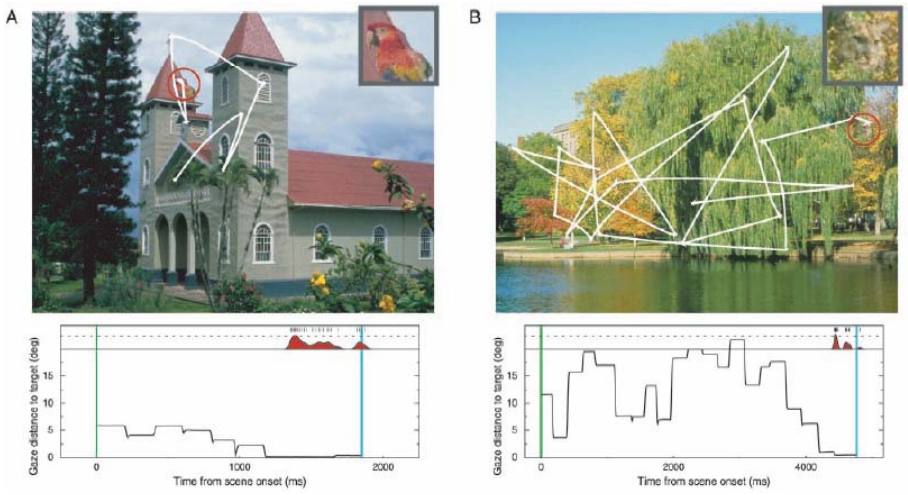
Sheinberg & Logothetis



1-12


A
Spikes/sec
150
100
50
0
0 150 300
S075
B
Spikes/sec
200
100
0
0 150 300
Time from stimulus onset (ms)
Q101

1-13


A
B
Gaze distance to target (deg)
15
10
5
0
0
1000
2000
Time from scene onset (ms)
Gaze distance to target (deg)
15
10
5
0
0
2000
4000
Time from scene onset (ms)

1-14