

Natural Scene Categorization: from Humans to Computers

Li Fei-Fei

Computer Science Dept.

Princeton University





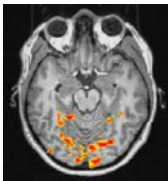
#1: natural scene categorization entails little attention



#2: what can we perceive within a glance of a scene?



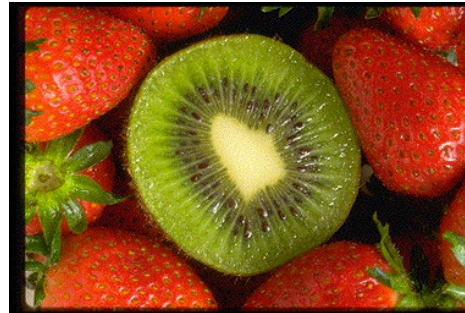
#3: Bayesian graphical models for natural scene categorization and event recognition



#4: decoding the neural representation of natural scene categories

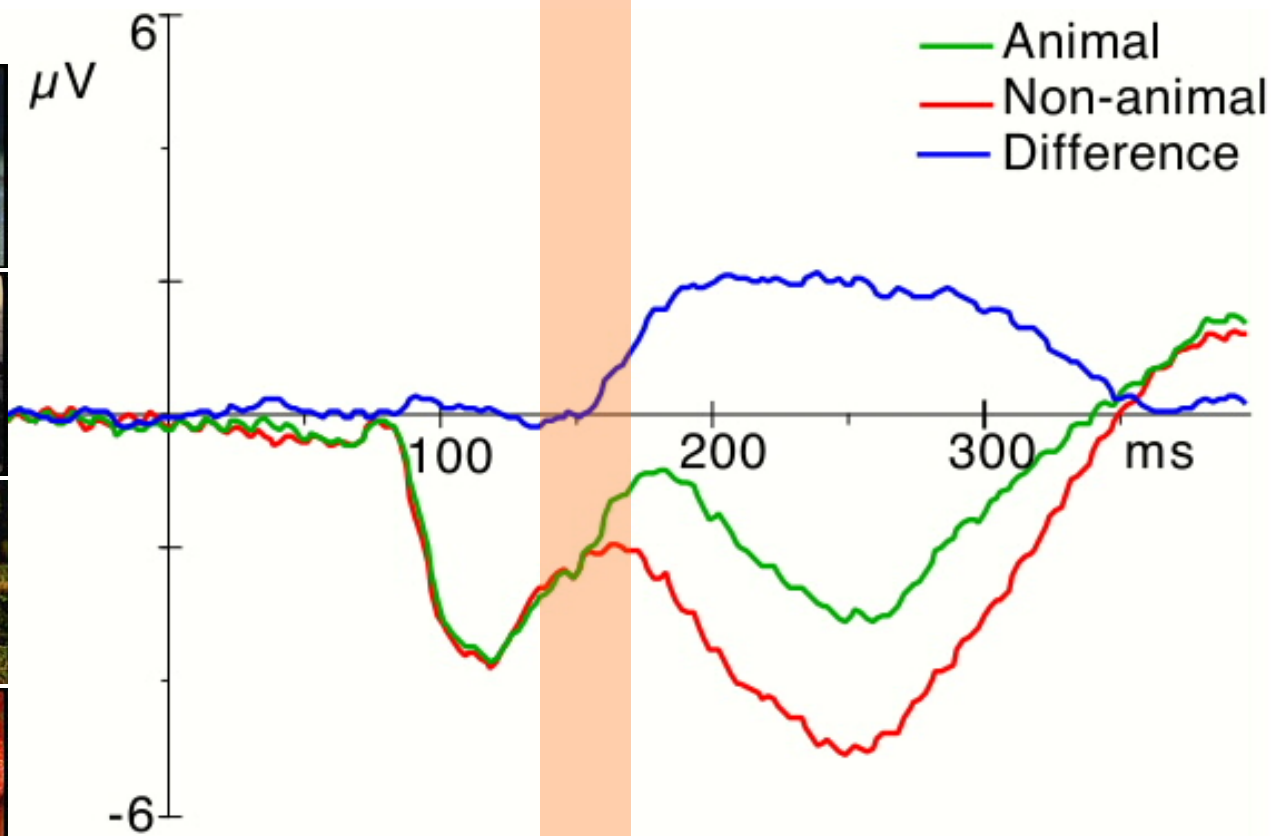


- #1: natural scene categorization entails little attention



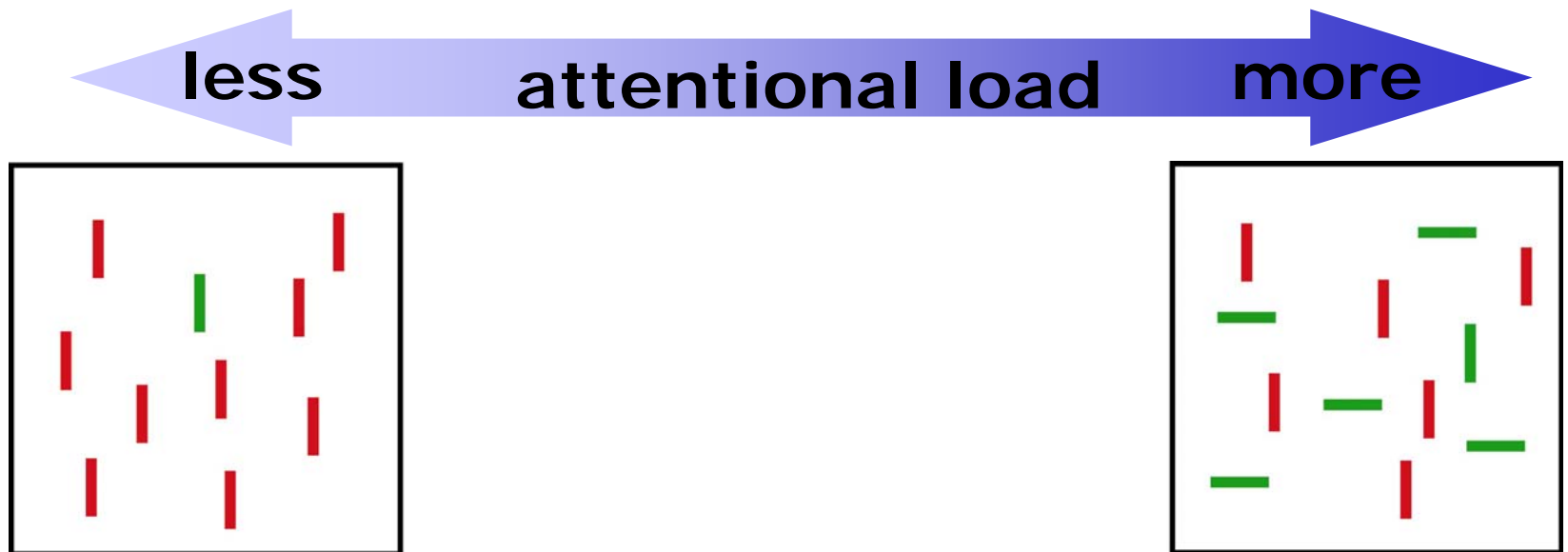
Reference: Li et al. 2002; Fei-Fei et al. 2005





150 ms !!

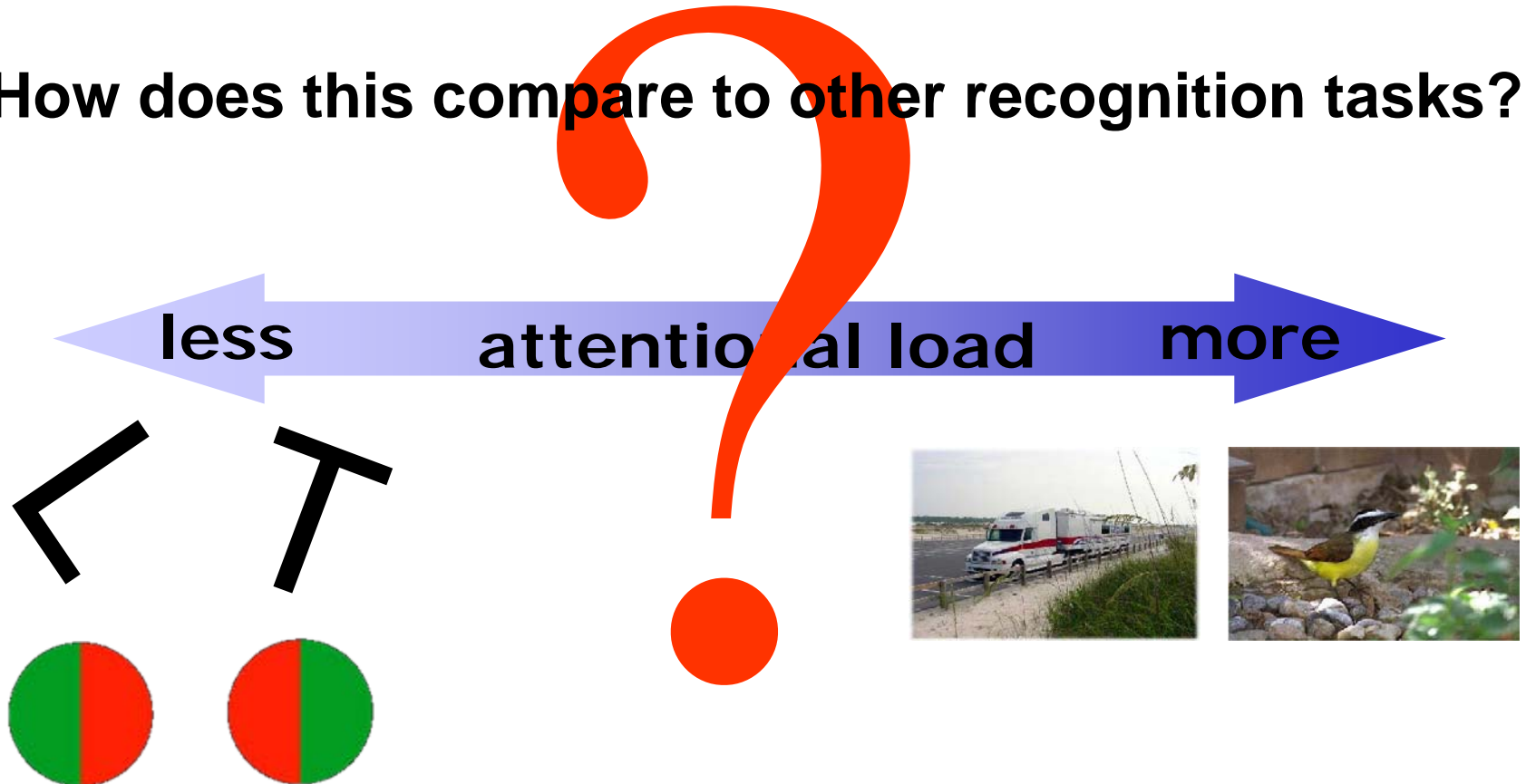
Attention binds features for recognition

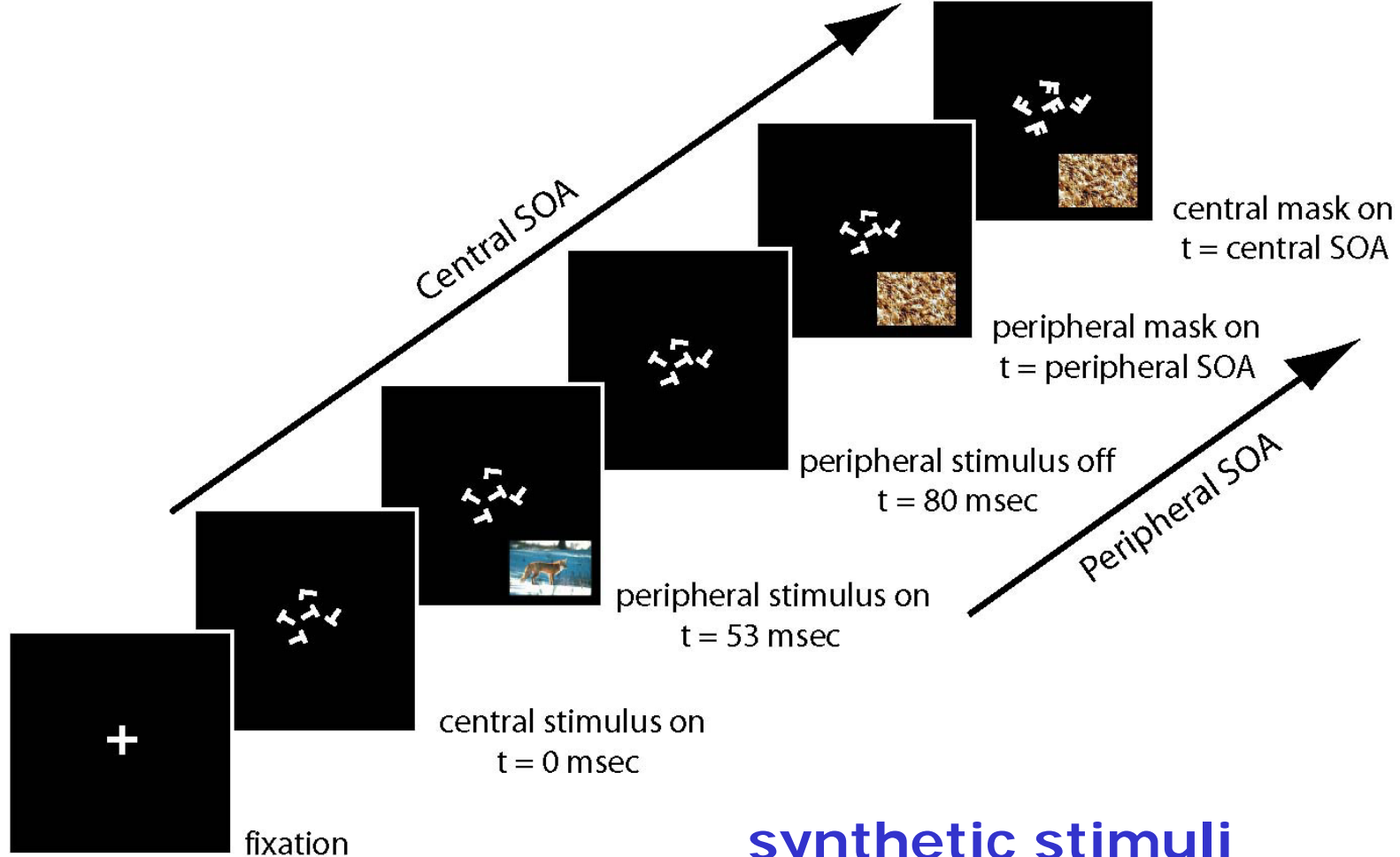


Our question

1. How critical is attention in natural scene recognition?

2. How does this compare to other recognition tasks?





animals



vehicle



synthetic stimuli



a Target: animal



vs.

Distractors



b Target: vehicle



vs.

Distractors



c Target: vehicle



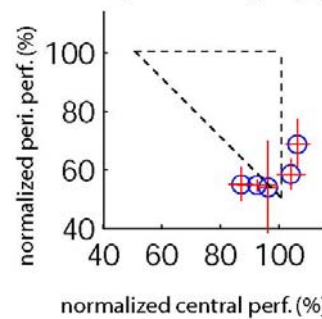
vs.

Distractors

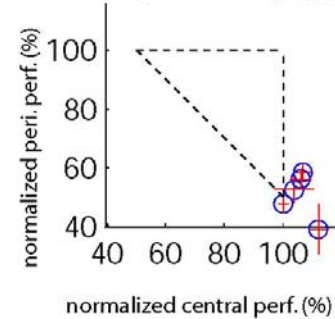


our finding...

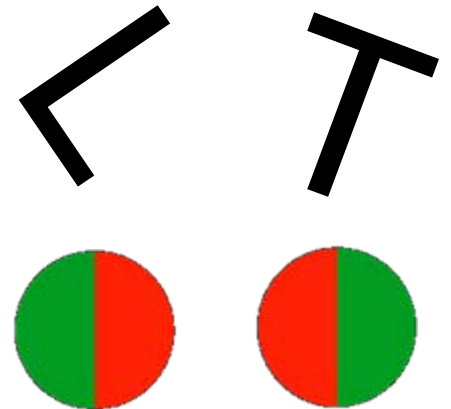
d  vs. 
(masked by )



e  vs. 
(masked by )



less **attentional load** more

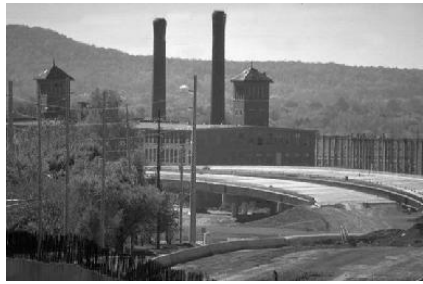


- #2: what can we perceive within a glance of a scene?

What do we
see in a glance
of a scene?

Reference: Fei-Fei et al. JoV 2007





Stage I: Collect Image Description

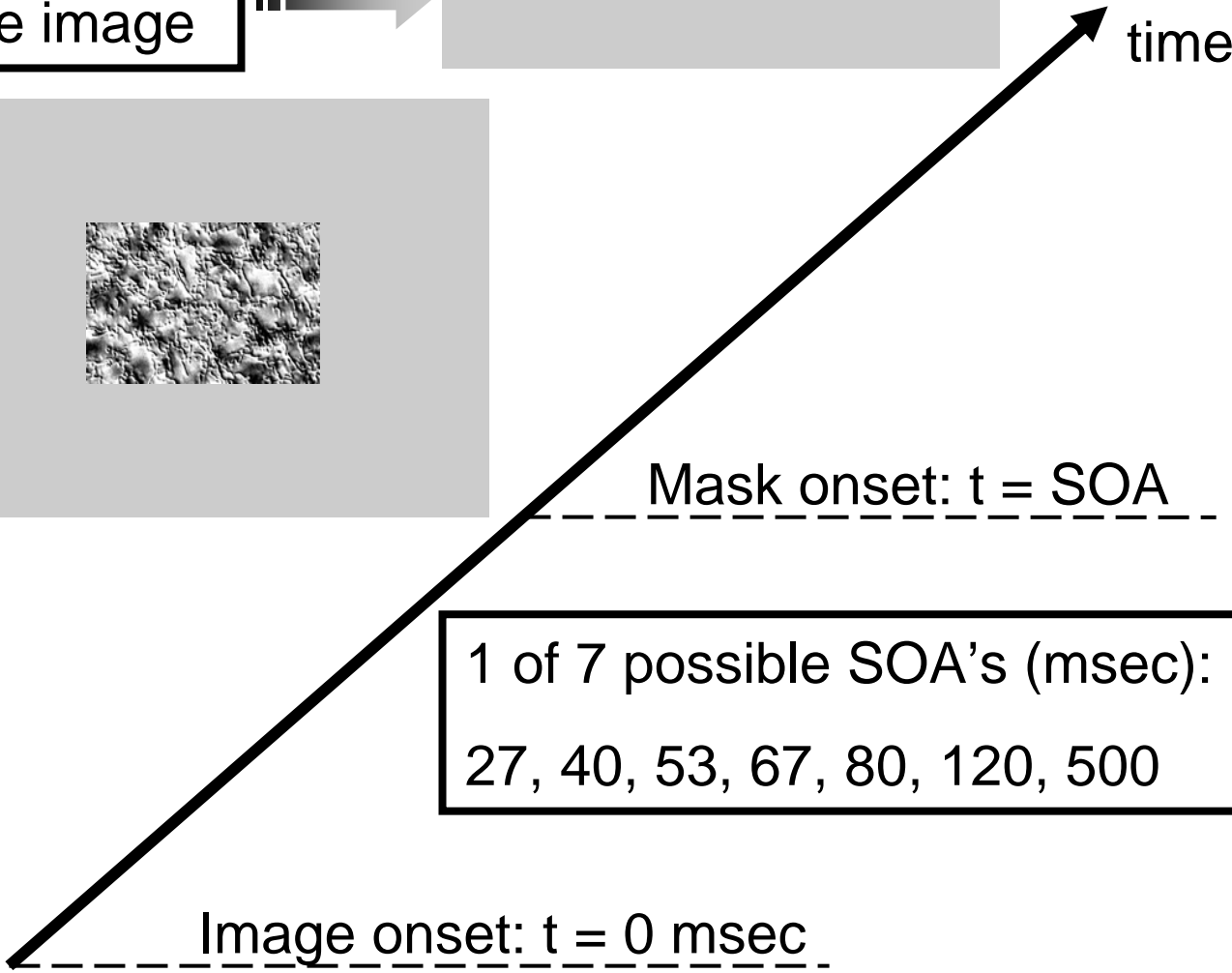
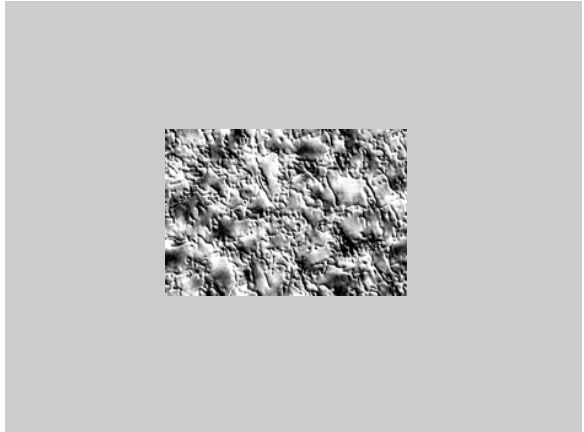
--- Illustration of 1 Trial

Subject types freely what he/she saw in the image



Please type your description here:

An outdoor scene, I think. reminded me a a city... like walking in a park in new york or something. there seemed to be trees and a road and then this large skyscraper in the background.



1 of 7 possible SOA's (msec):
27, 40, 53, 67, 80, 120, 500



PT = 27ms

Couldn't see much; it was mostly dark w/ some square things, maybe furniture. (Subject: AM)

PT = 40ms

This looked like an indoor shot. Saw what looked like a table and chairs.

22 subjects, 100 images, 2200 responses

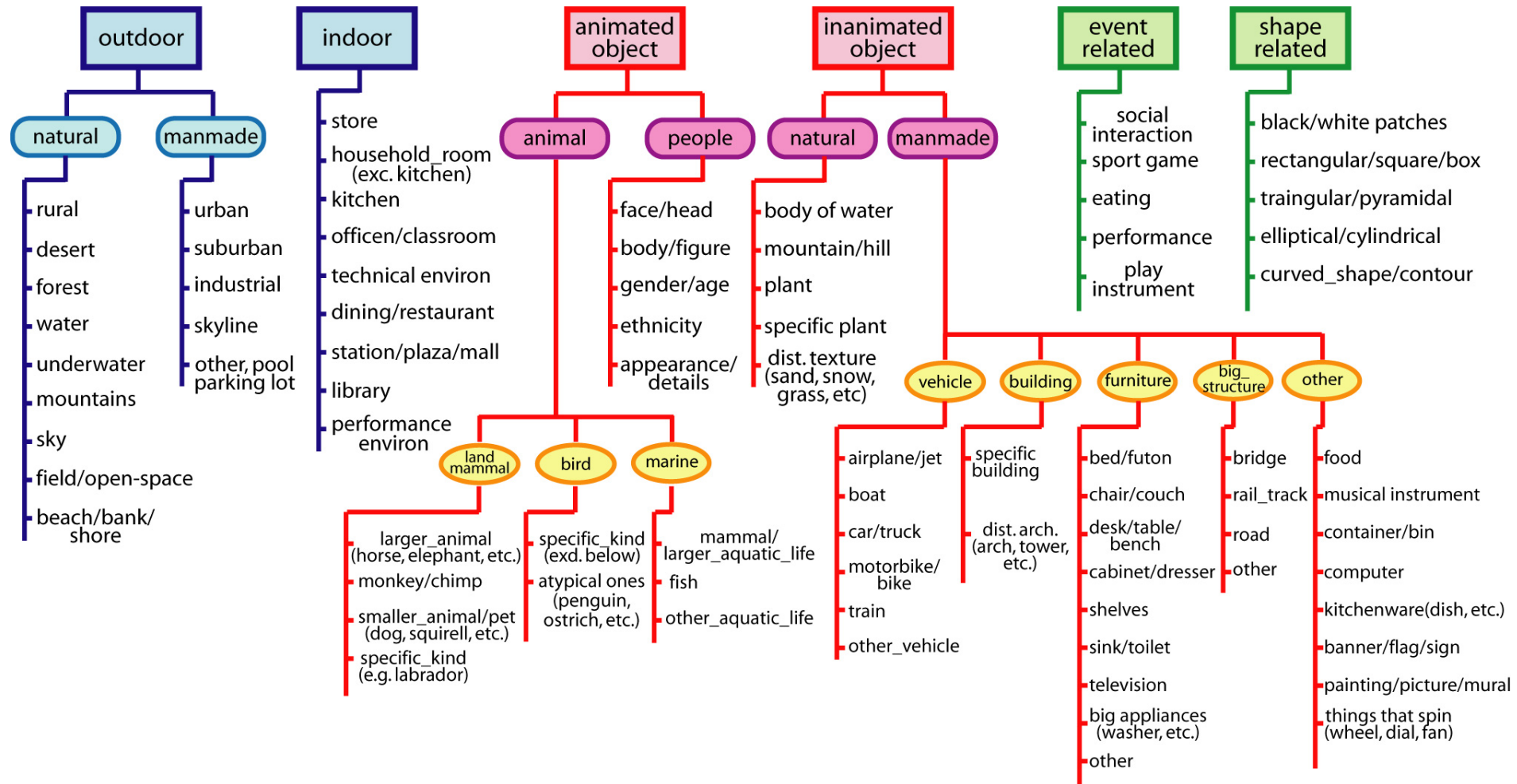
There are many paintings on the wall. The largest painting might have a fireplace beneath it. I think the largest painting was that of a man standing erect. The room is richly decorated and it looks like one of the rooms in Mr. Darcy's house in the A&E movie *Pride and Prejudice*. Or maybe it more closely resembles one of the rooms where the one of the rooms in Huntington's house (at the Huntington).

... (Subject: RW)

PT = 67ms

I saw the interior of a room in a house. There was a picture to the right, that was black, and possibly a table in the center. It seemed like a formal dining room. (Subject: JB)

The attribute tree



Response No. 18 for Image No. 4



CATEGORY: SENSORY/SHAPES

Please select one of "correct" or "incorrect" for each checked description. Click "Next>>" to continue

- black/white_patches described correct incorrect
- rectangular/square/box described correct incorrect
- triangular/pyramidal described correct incorrect
- elliptical/cylindrical(eg.round,blob) described correct incorrect
- curved_shape/contour(eg.arc,'S') described correct incorrect

I could make out some kind of circular shapes near the bottom of the picture. These reminded me of those round life preservers that are on ships. There was also a man standing on top of some wooden structure.

What's in a single glance?



PT = 107ms

This is outdoors. A black, furry dog is running/walking towards the right of the picture. His tail is in the air and his mouth is open. Either he had a ball in his mouth or he was chasing after a ball. (Subject EC)

PT = 500ms

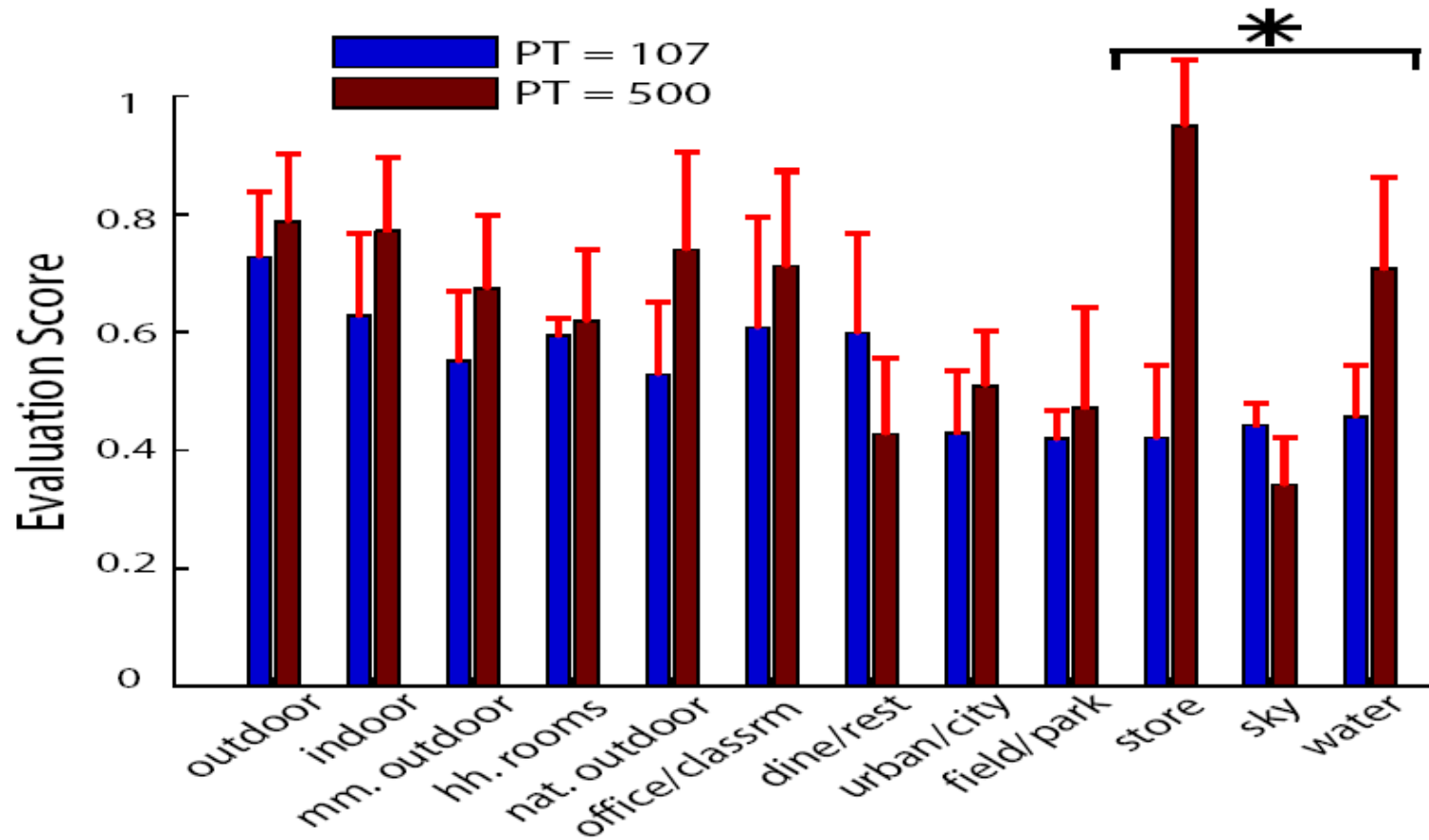
I saw a black dog carrying a gray frisbee in the center of the photograph. The dog was walking near the ocean, with waves lapping up on the shore. It seemed to be a gray day out. (Subject JB)



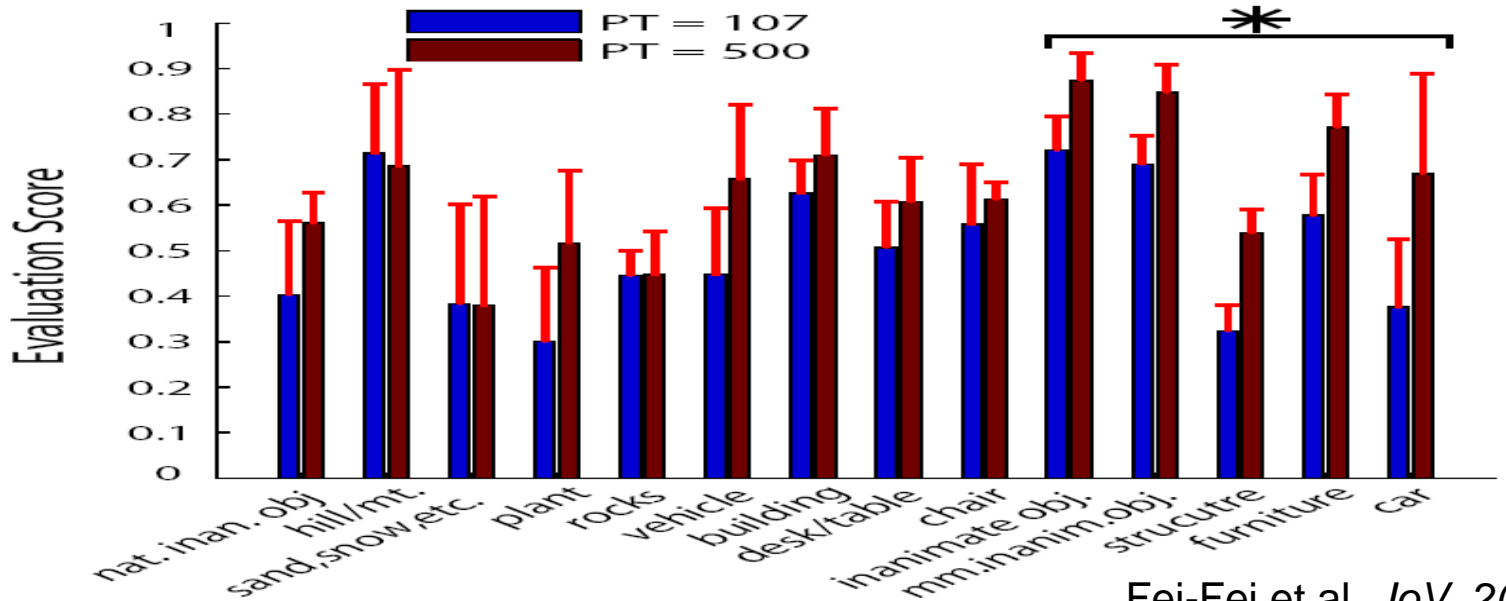
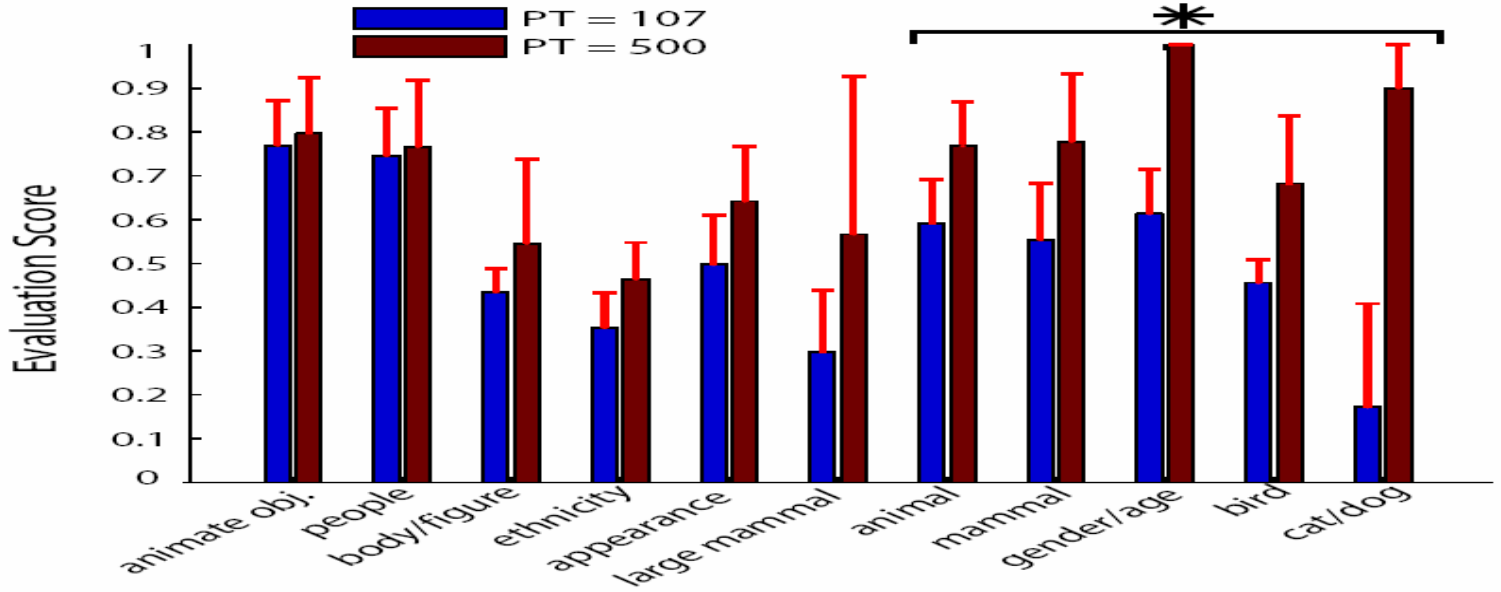
inside a house, like a living room, with chairs and sofas and tables, no ppl. (Subject HS)

A room full of musical instruments. A piano in the foreground, a harp behind that, a guitar hanging on the wall (to the right). It looked like there was also a window behind the harp, and perhaps a bookcase on the left. (Subject RW)

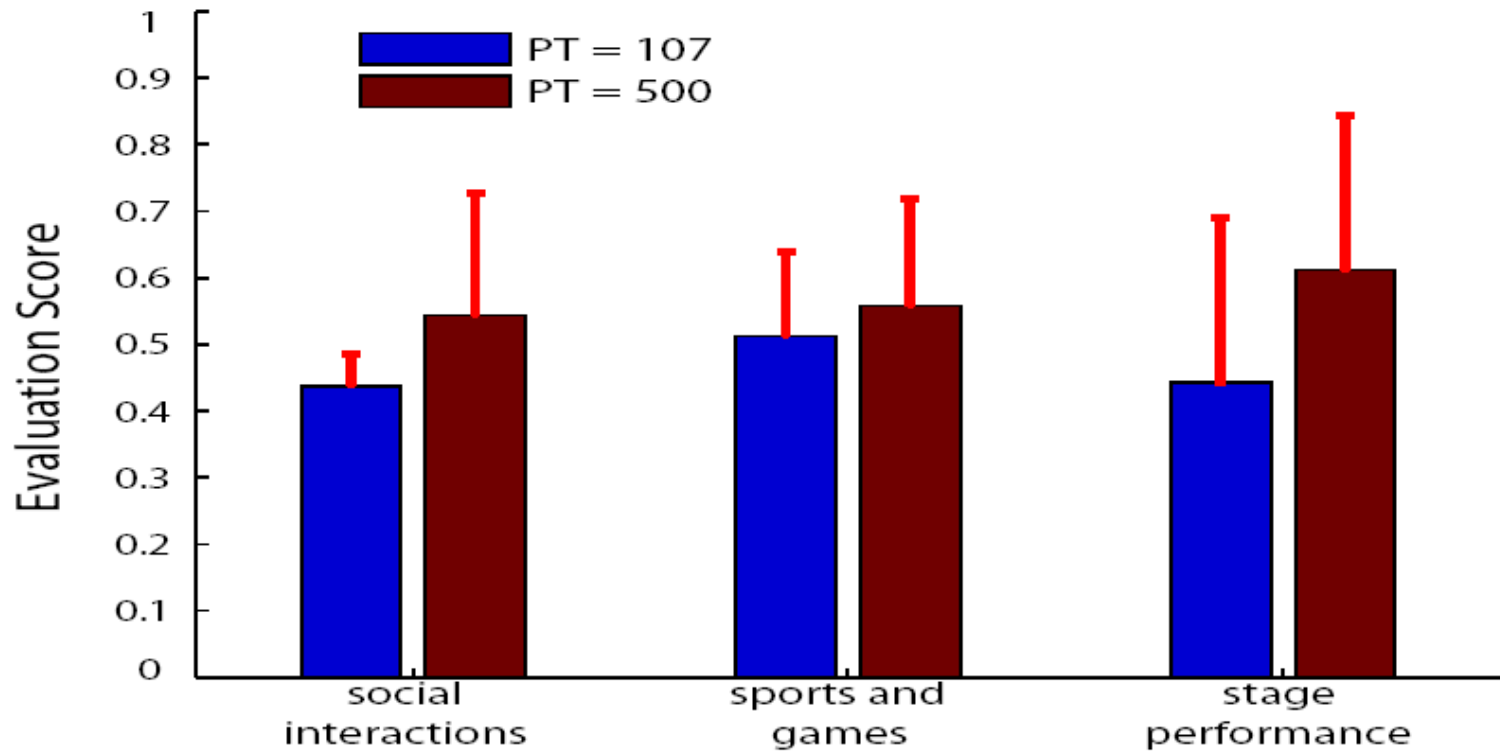
Scene level



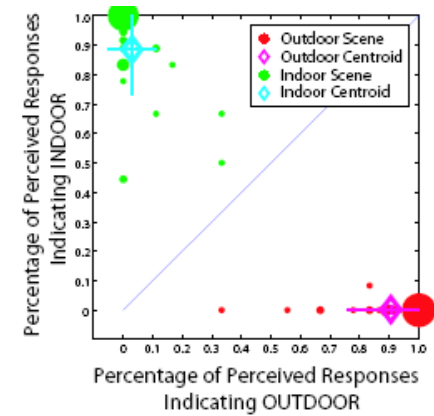
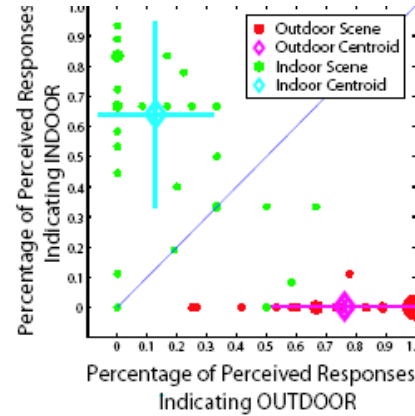
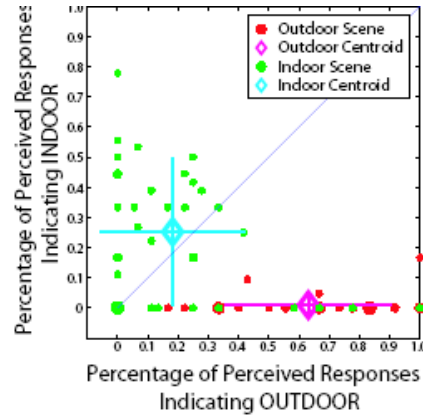
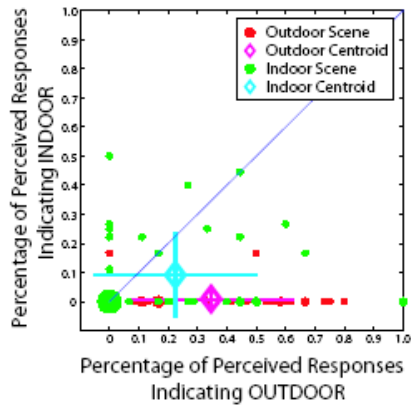
Object level



(Social) Events



Indoor vs. outdoor categorization



(a) PT 40ms

(b) PT 67ms

(c) PT 107ms

(d) PT 500ms

from human vision we know...

beach



living room

city



**natural scene
categorization**

model representation

Image

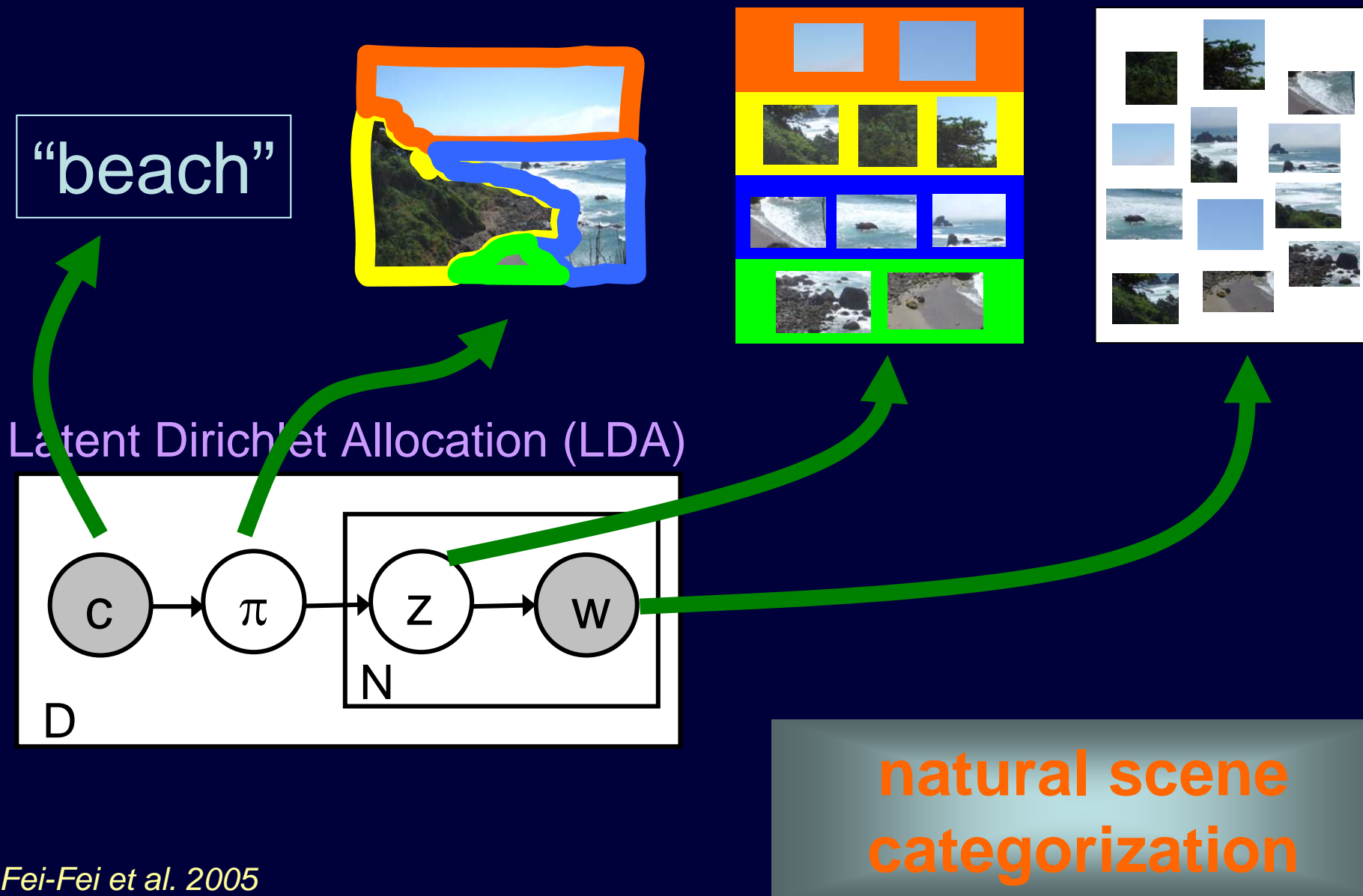


Bag of 'visual words'

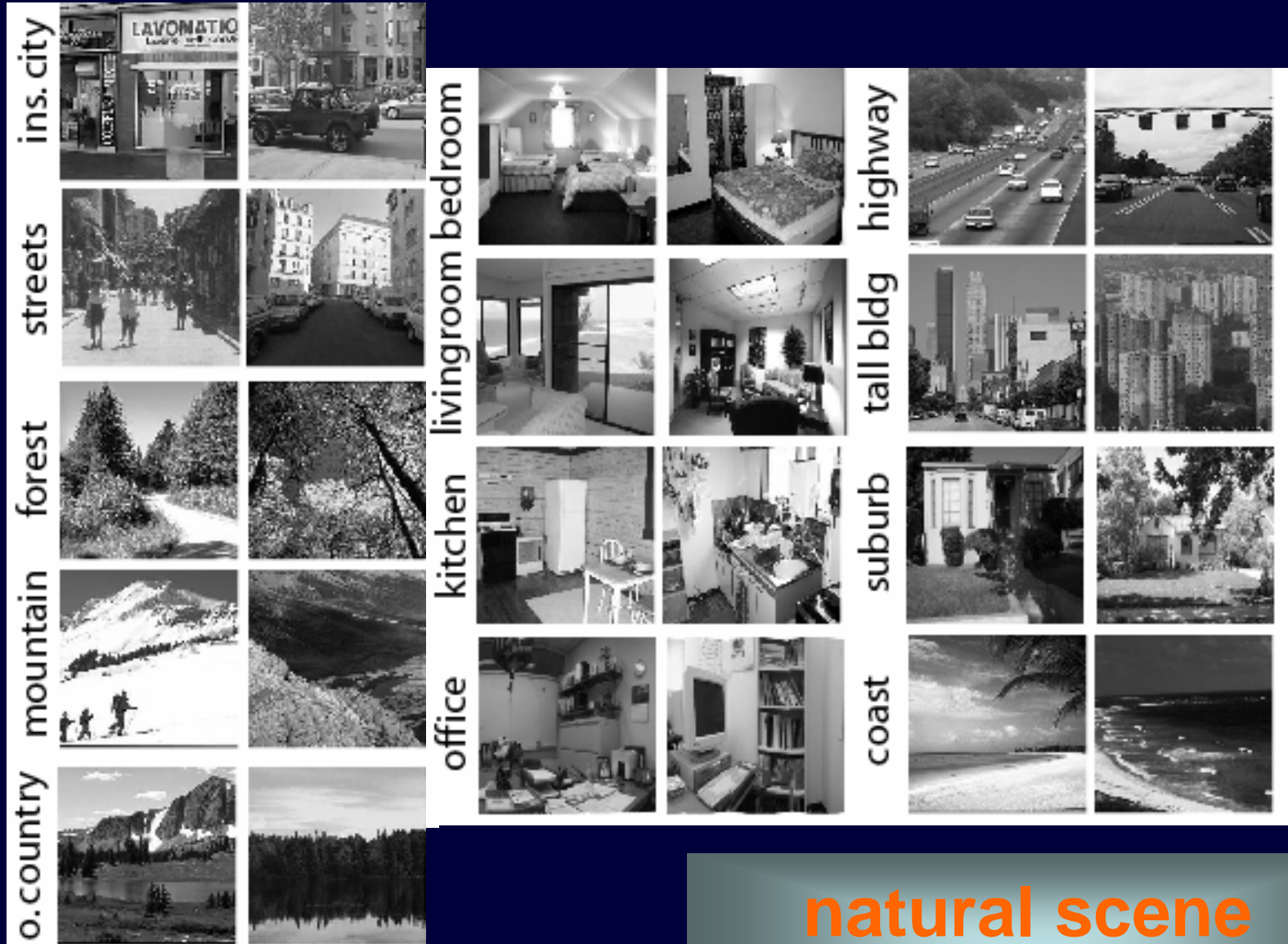


natural scene
categorization

model representation & learning

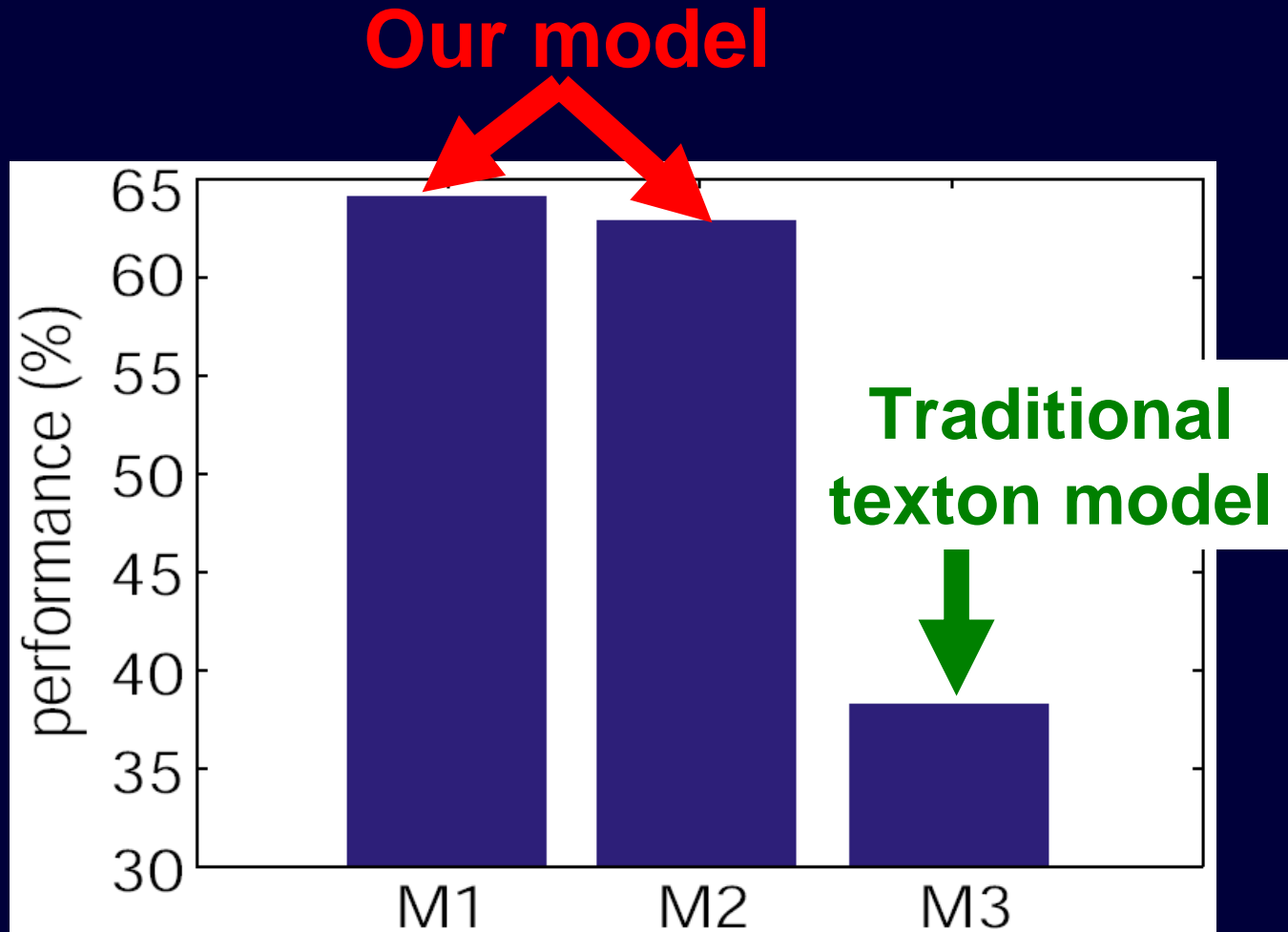


evaluation & dataset



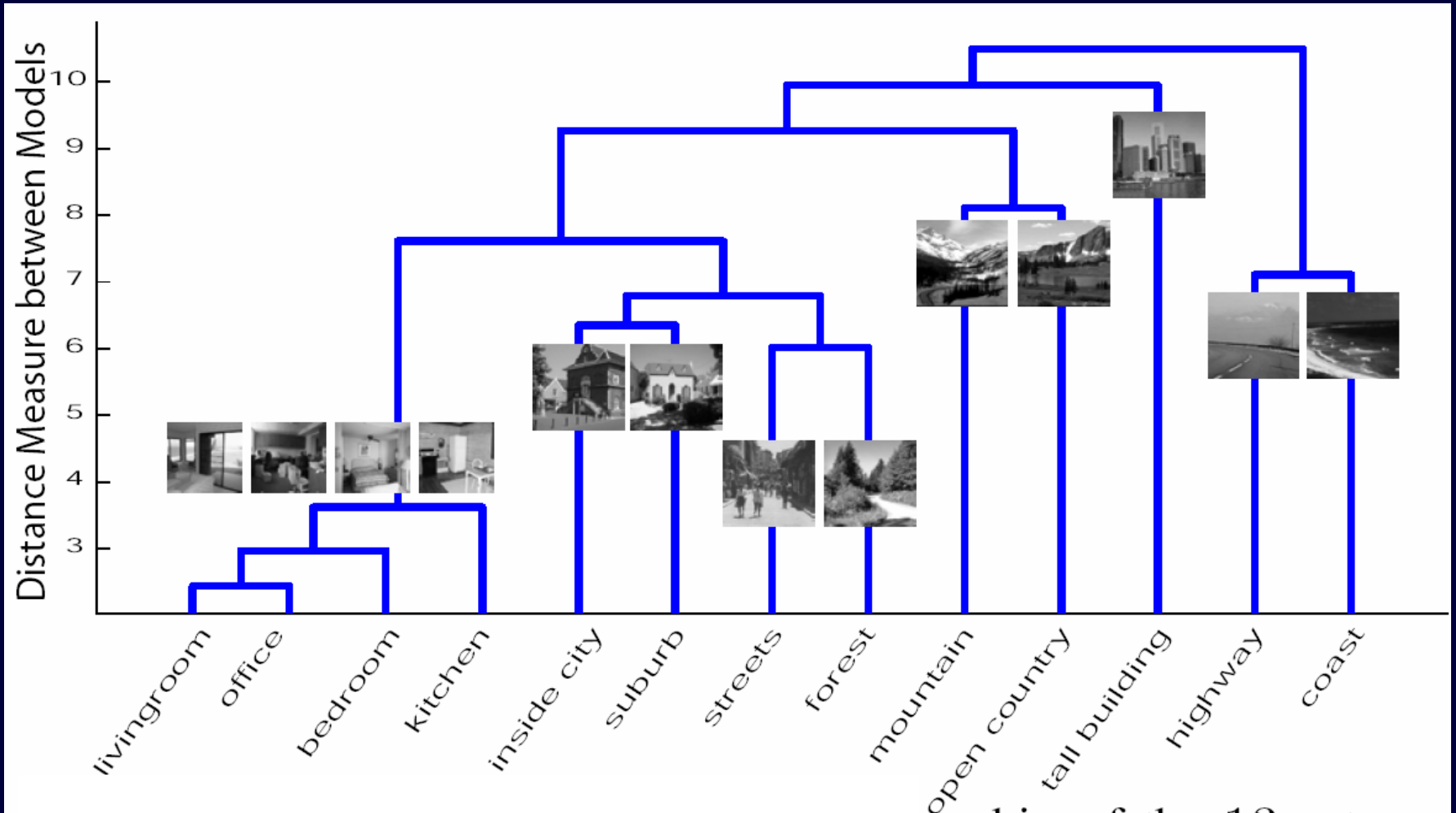
**natural scene
categorization**

evaluation & dataset



**natural scene
categorization**

evaluation & dataset



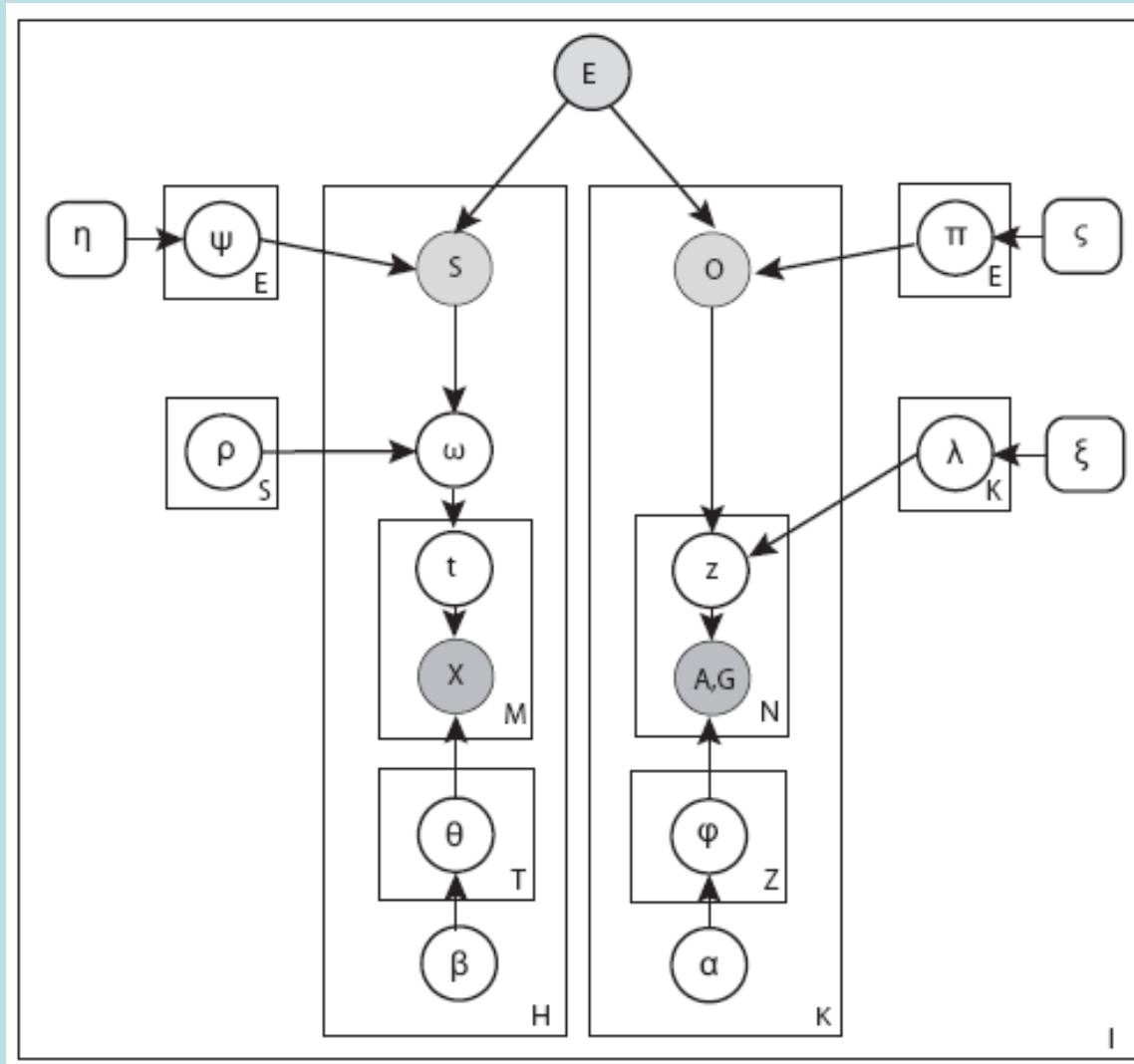
**natural scene
categorization**



PT = 500ms

This is indoors. It's must be a rich person's house. There are many paintings on the wall. The largest painting might have a fireplace beneath it. I think the largest painting was that of a man standing erect. The room is richly decorated and it looks like one of the rooms in Mr. Darcy's house in the A&E movie *Pride and Prejudice*. Or maybe it more closely resembles one of the rooms where the one of the rooms in Huntington's house (at the Huntington).

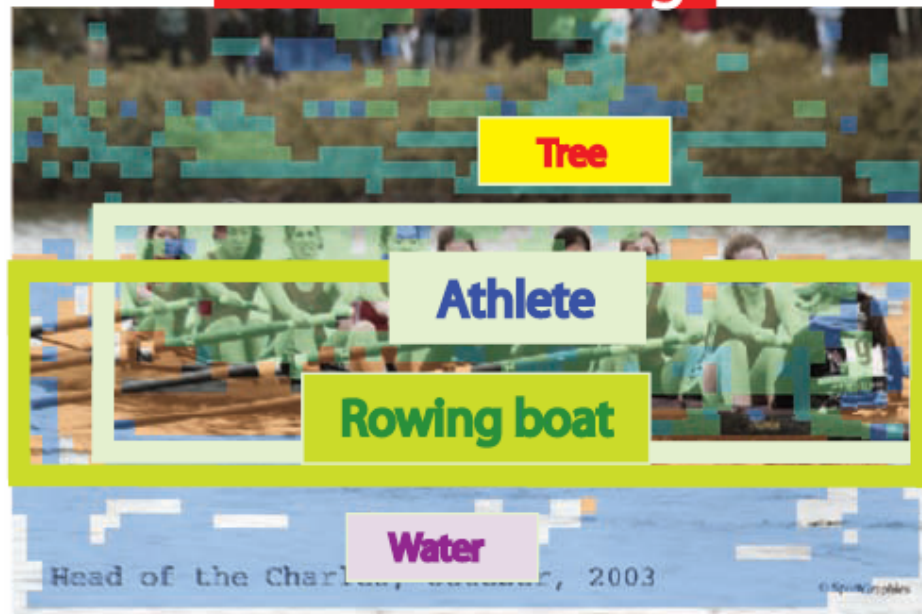
Event recognition



Event recognition



event: Rowing



scene: Lake

Event recognition

event: Badminton



scene: Badminton court

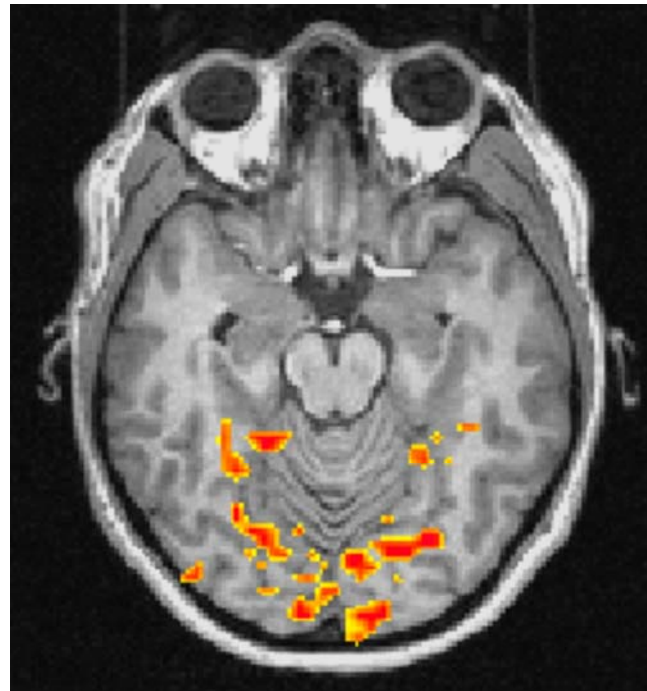
event: Rockclimbing



scene: Mountain

Event recognition

- #4: decoding the neural representation of natural scene categories



Reference: Et al. Fei-Fei, VSS 2007, HBM 2007



Highways



Mountains



Buildings



Industry

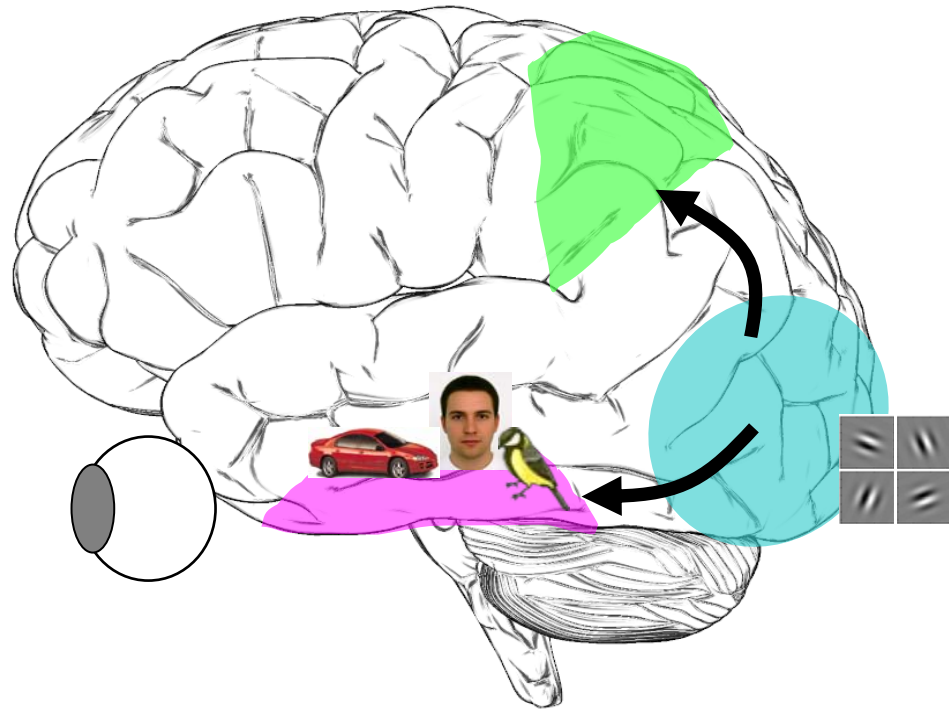
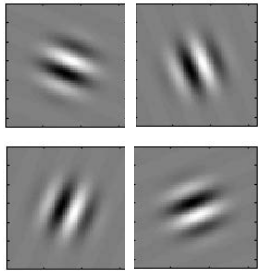


Beaches



Forests





Hubel & Wiesel 1961, Kanwisher et al. 1997

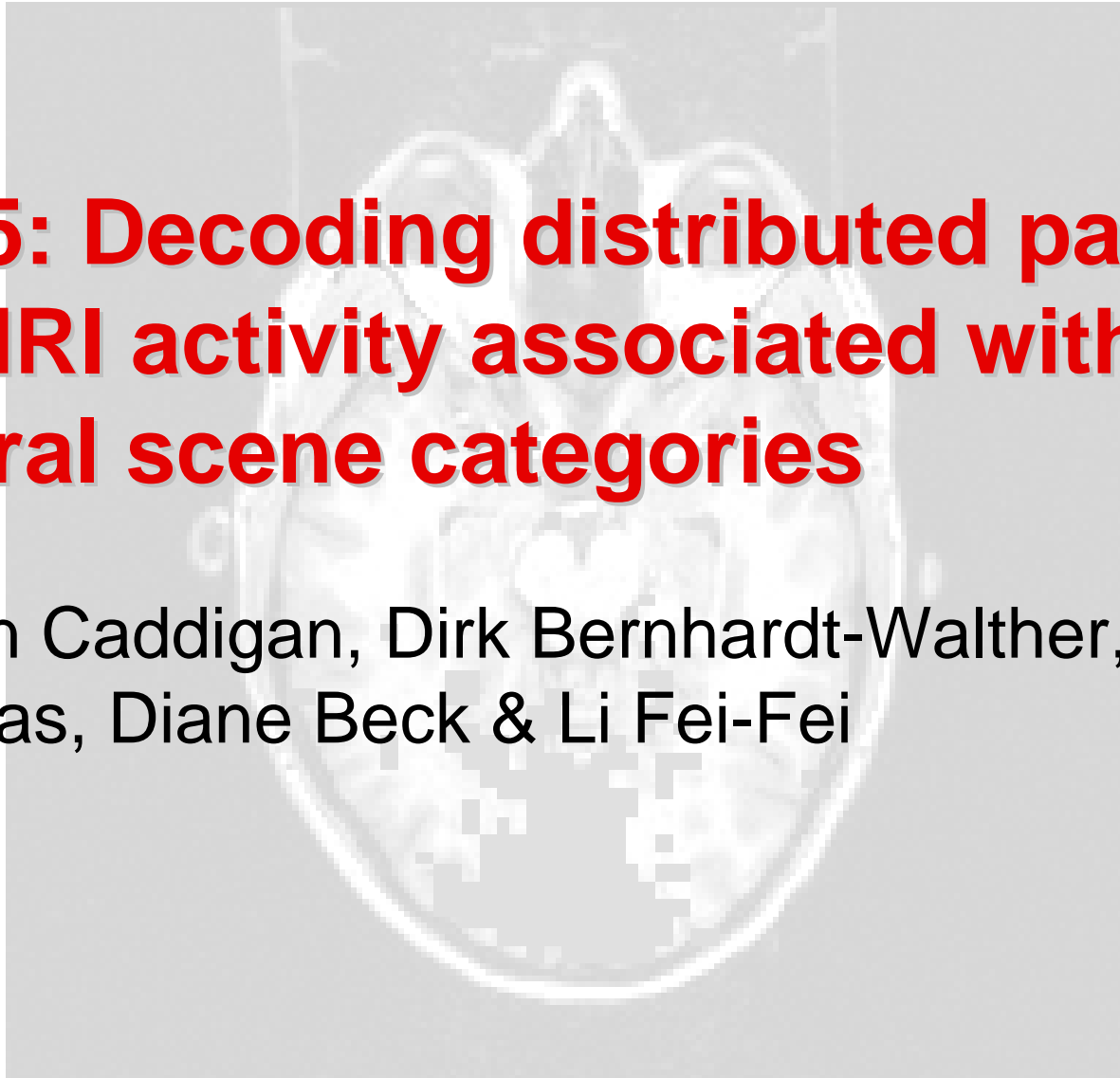


Epstein & Kanwisher, 1998

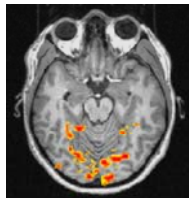
Tuesday May 15, Talk Session 9:00am

**# 865: Decoding distributed patterns
of fMRI activity associated with
natural scene categories**

Eamon Caddigan, Dirk Bernhardt-Walther, Justas
Birgiolas, Diane Beck & Li Fei-Fei



Summary



- natural scene categorization entails little attention
(Rufin VanRullen, Pietro Perona, Christof Koch)
- natural scene at a glance: ‘gist’ of a scene includes much information on objects, scenes and beyond
(Asha Iyer, Pietro Perona, Christof Koch)
- hierarchical graphical models for natural scene categorization and event recognition
(Li-Jia Li)
- decoding the neural representation of natural scene categories from patterns associated with fMRI activity
(Diane Beck, Eamon Caddigan, Dirk Walther)



Thank you!

