

Looking at Things and Stuff

SunS 2009

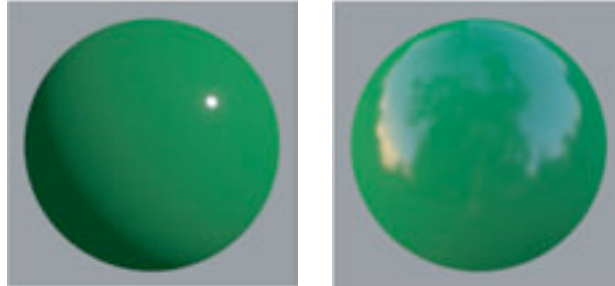
Edward H. Adelson
(with **Lavanya Sharan**
& Ruth Rosenholtz)

We can see objects, scenes, and the “stuff” (materials) they are made of.



- People can describe material attributes like color, translucency, etc. I.e., they can estimate parameters.
- People can name materials (plastic, wood, etc.) I.e., they can recognize categories.
- Most material research is about parameter estimation, not recognition/categorization.

Parameter estimation research uses simple toy worlds.



*Rendered spheres (Glossiness)
Fleming, Dror & Adelson 2003*



*Photographs of stucco (Albedo
& Gloss) Motoyoshi et al 2007,
Sharan et al 2008*

Judgments like albedo, gloss, color, translucency etc. have been studied.

Beck & Prazdny 1981, Adelson 1993, 1999; Nishida & Shinya 1998, Bloj et al. 1999, Gilchrist et al. 1999, Pellacini et al. 2000, Singh & Anderson 2002, Boyaci et al. 2003, Brainard et al. 2003, Fleming et al. 2003, Maloney & Yang 2003, Fleming et al. 2004, Todd et al. 2004, Berzhanskaya et al. 2005, Fleming & Buelthoff 2005, Khang et al. 2006, Robilotto & Zaidi 2006, Ho et al. 2007, Motoyoshi et al. 2007, Vangorp et al. 2007, Sharan et al. 2008, Xiao & Brainard 2008.

Let's study recognition in natural images.
It can be very fast.

- Potter 1976 (gist in RSVP)
- Thorpe et al (many) (is there an animal here?)
- Greene & Oliva 2009 (what kind of scene is this?)
- Fei Fei et al 2007 (describe this picture)
- Grill-Spector & Kanwisher 2005 (what category object?)

How about material perception?

(Sometimes it's slow, e.g., deciding if leather is real or fake.
But, we find, sometimes it can be quite fast.)

What do people mean by “rapid perception?”

- You can see a lot with a short presentation. (But maybe processing continues for a long time)
- You can make judgments with short RT.

Start with simple experiment: show a picture briefly, get description (cf. Fei Fei et al)

Subject's description after 320msec glimpse:

“A wool cupcake, a knitted wool cupcake, so the cupcake part was made out of light brown wool with small specks of yellow and maybe pink wool on top, so looked like sprinkles with a red wool cherry...”

Subject sees the thing (cupcake) and the stuff (knitting) quickly and separately.

Object and material perception are distinct.



Can you convey material info verbally?

Make all objects have same shape, so shape is uninformative.

- Subject A views a “donut” shaped object for 40 msec or 320 msec, followed by mask.
- Subject A then describes the materials in words (no color words or object names allowed)
- Subject B reads the description, tries to match it with the correct picture.



Subject A describes materials.
Subject B matches to images.

60 pictures



Etc.

60 descriptions

→ “Blah blah...”

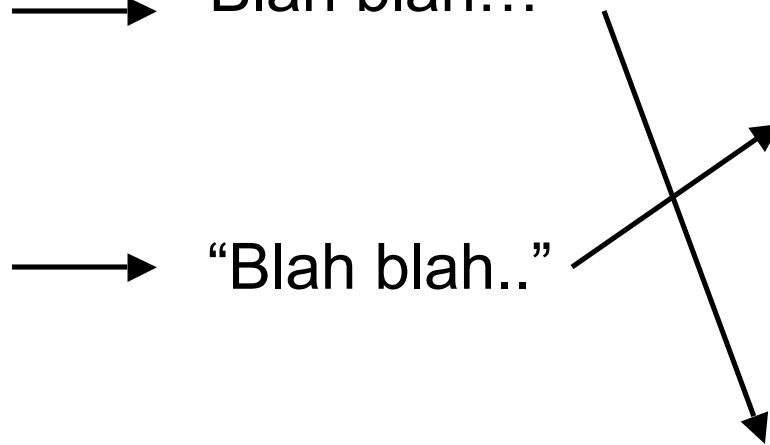
→ “Blah blah..”

Etc.

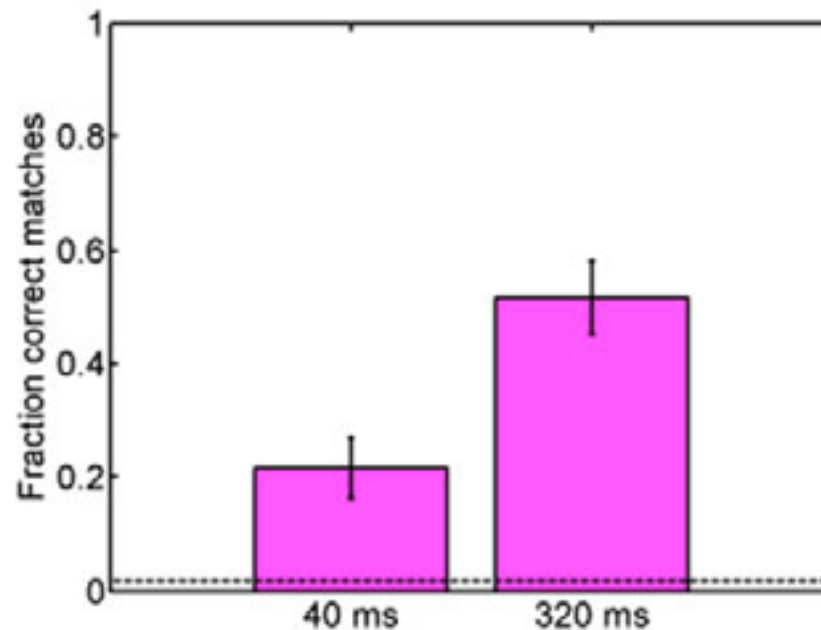
60 pictures



Etc.



Matching is above chance at 40 msec,
improves at 320 msec.



Longer glimpses are better, but even short glimpses provide useful information.

Let's do material categorization. Build database with nine categories:

Fabric, glass, leather, metal, paper, plastic, stone, water, wood.



Our images are diverse. (Here, fabric)



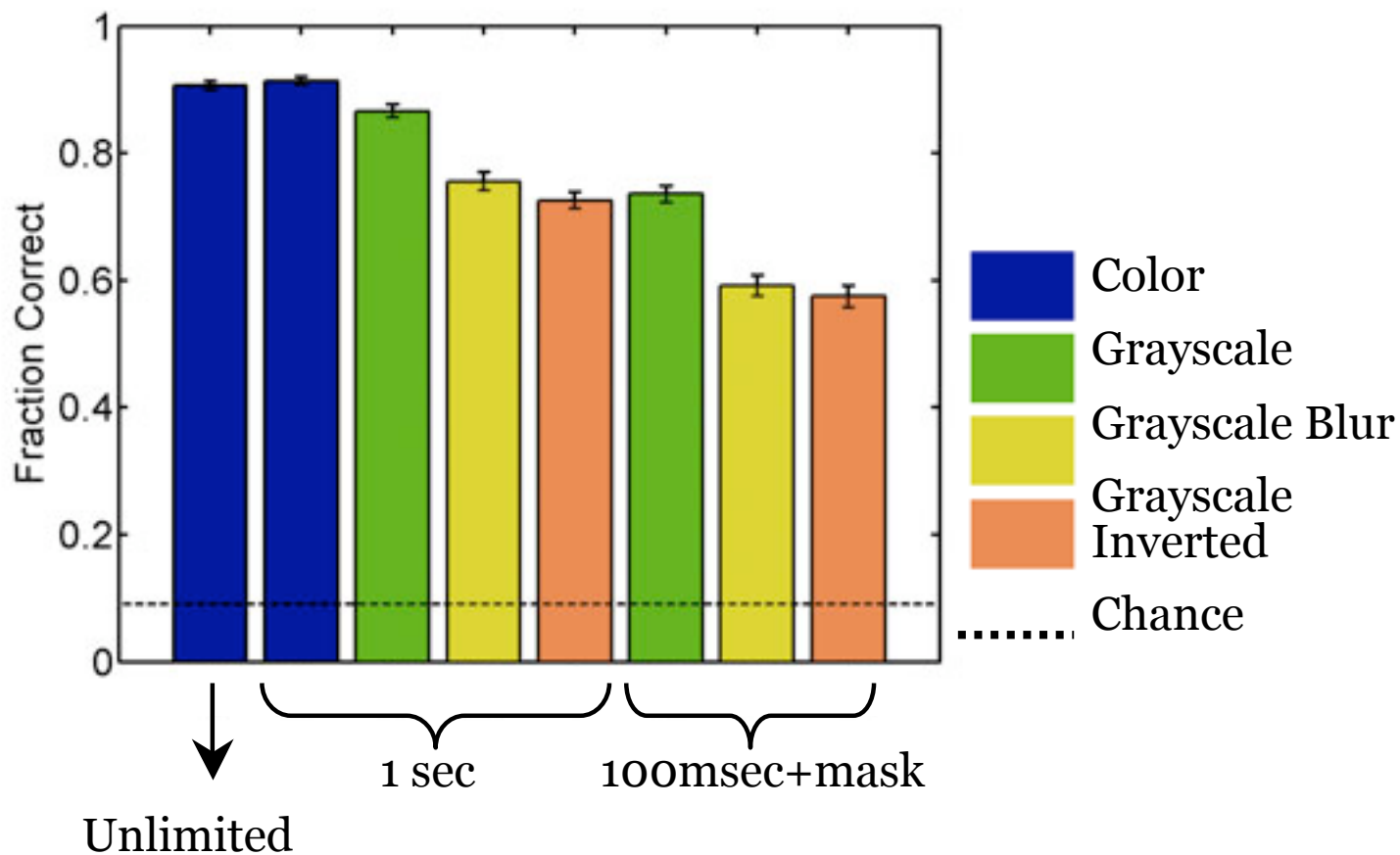
Our images are diverse. (Here, plastic)



Image diversity means no simple
“tricks” let you do the task.

- You can't just use low-level features like color and texture.
- You can't just do object recognition.

Categorization (9 categories) of single images, under various conditions.



Surprisingly robust!

Make it really hard: RSVP stream

40 or 160 ms
per image

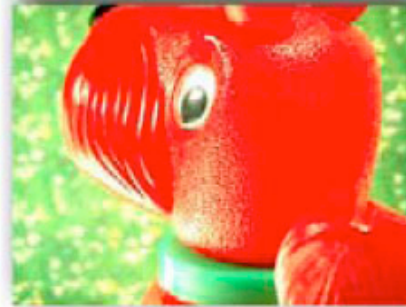


Position 1



Position 2

...

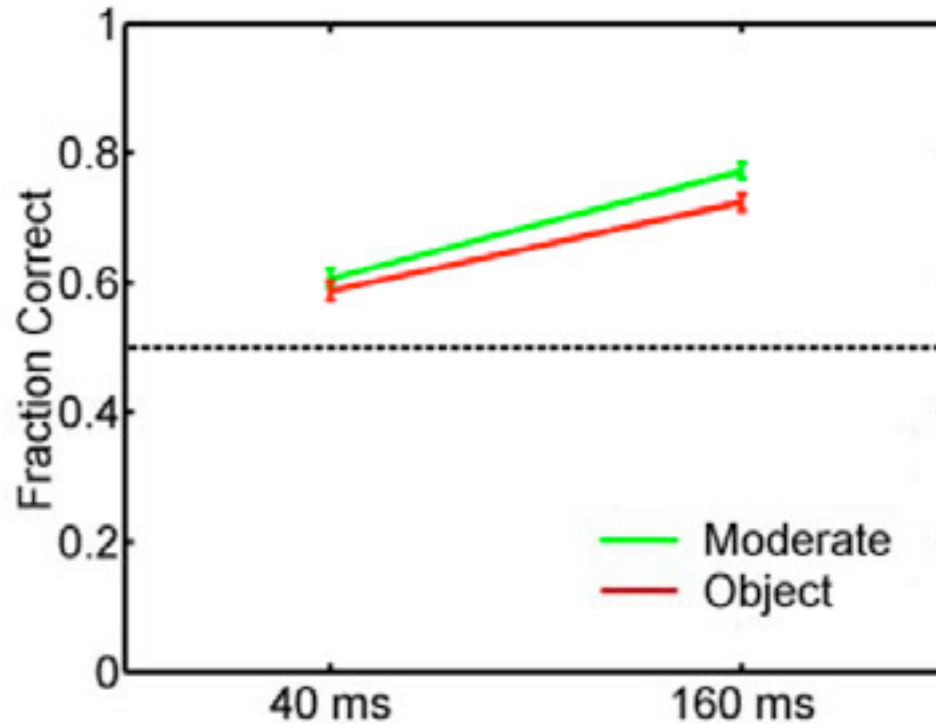


Position 9

Target appears at Positions 3-7

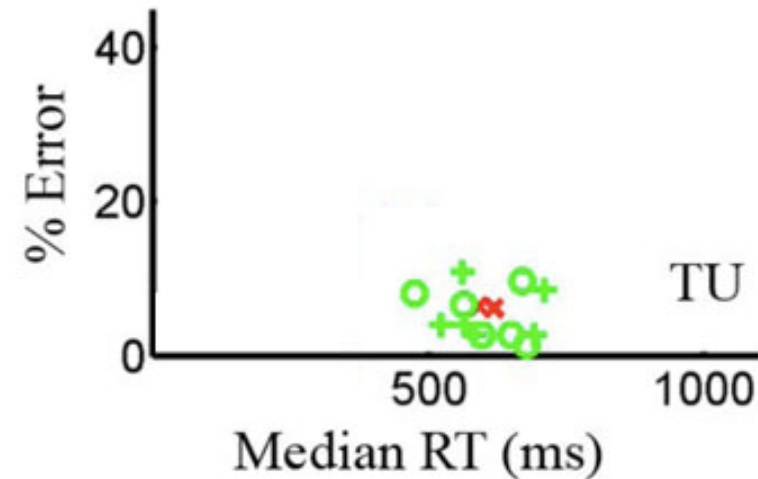
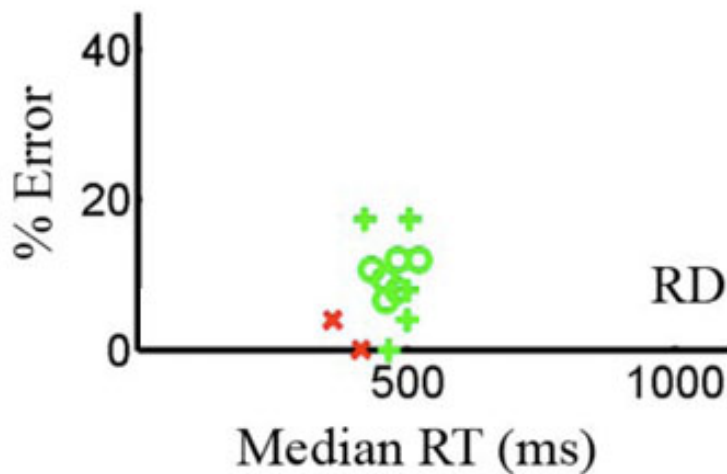
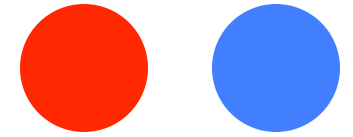
Task: is there a picture of (e.g.) metal in the stream?

- **Forced choice. Chance = 50%**
- **9 images. Target only in positions 3-7.**
- **It's tough! But people are above chance.**



Plain old reaction time. Almost as fast as baseline!

- Display a picture of paper, plastic, or fabric.
- Subject chooses category as fast as possible.
- For baseline RT, use categorization of red versus blue disk



- Baseline task
 - Categorization task
- (+ = BW, o = color)

Conclusions

- Material perception is a distinct problem (not same as texture, object, scene recognition)
- Recognition can be fast and accurate in challenging conditions.
- We don't know how it works.