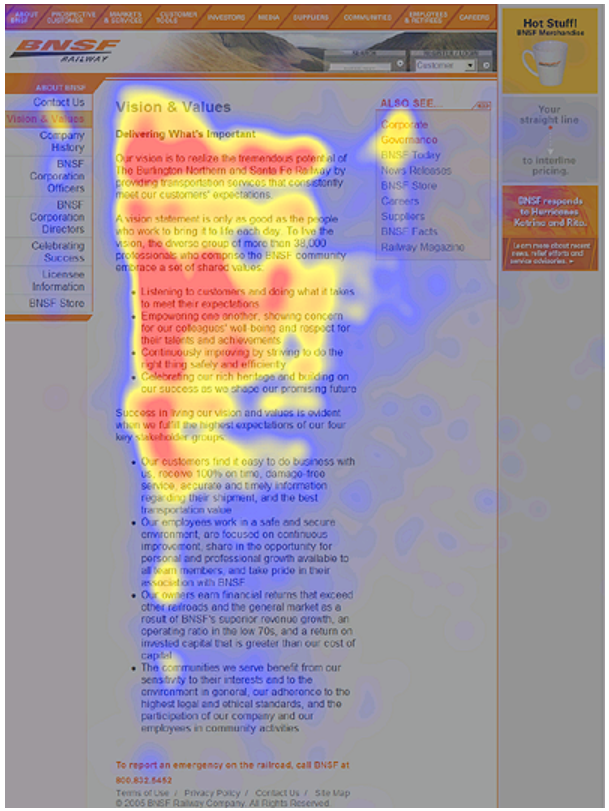
Using machine learning to predict where people look

Tilke Judd, Antonio Torralba, Frédo Durand

Understand where humans look

- For applications in graphics, design, robotics
- reading, skimming, search
- measure and reduce clutter



http://www.useit.com/alertbox/reading_pattern.html

Applications

Doug DeCarlo and Anthony Santella [SIGGRAPH 2002]

Level of detail for NPR

(a) original (b) gaze-based (c) distracting crop

Santella et al. [SIGCHI 2006]

Automatic cropping

provide equal or improved performance without eye tracking equipment

Saliency

- measure of conspicuity
- likelihood of a location to attract attention of human

Saliency Models

(a) Original image

(b) Hou and Zhang

(c) Itti and Koch

- Itti and Koch

L. Itti and C. Koch. A saliency-based search mechanism for overt and covert shifts of visual attention, 2000.

- Ruth Rosenholtz

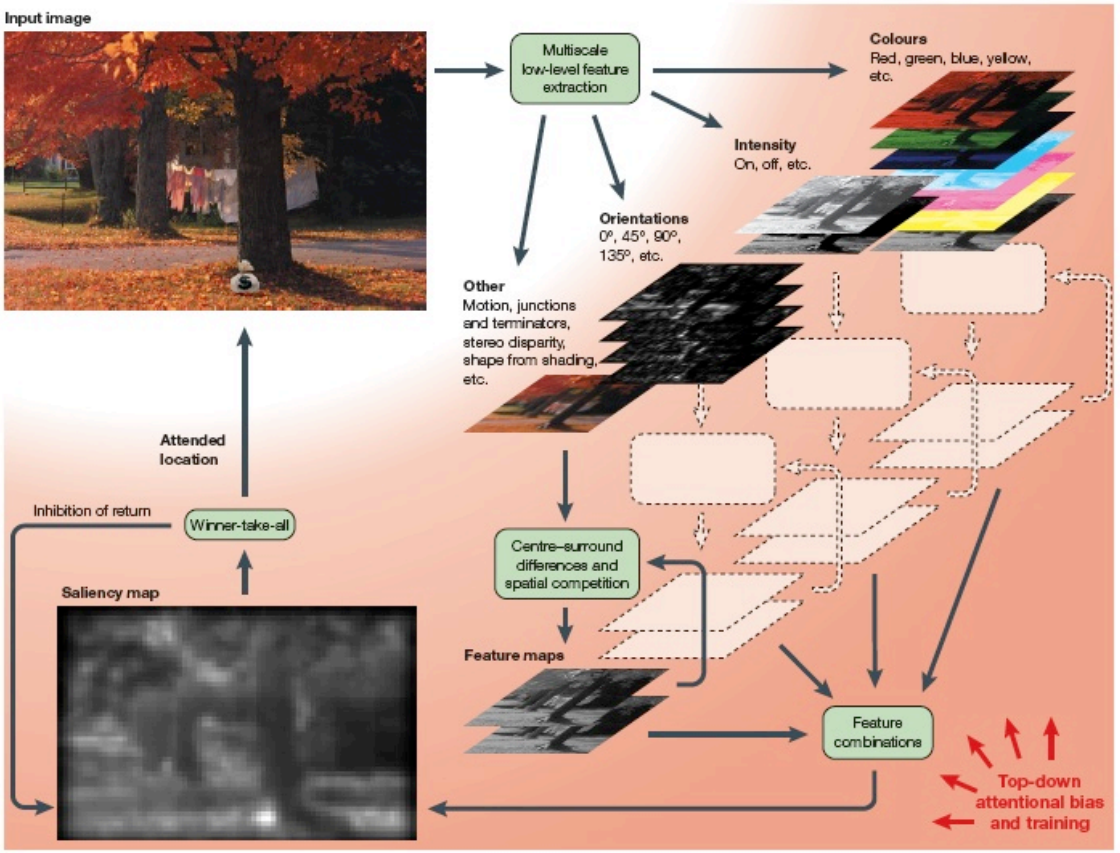
R. Rosenholtz. A simple saliency model predicts a number of motion popout phenomena. Vision Research 39, 19:3157–3163, 1999.

- Hou and Zhang

X. Hou and L.Q. Zhang. Saliency detection: A spectral residual approach. 2007.

Saliency Models

- based on biologically plausible linear filters
- measure intensity, illumination, and color contrast
- lots of parameters
- bottom up model

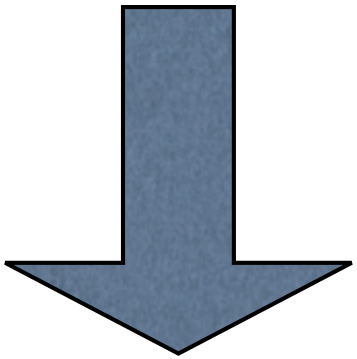


Bottom-up saliency model. From Itti and Koch [2001]

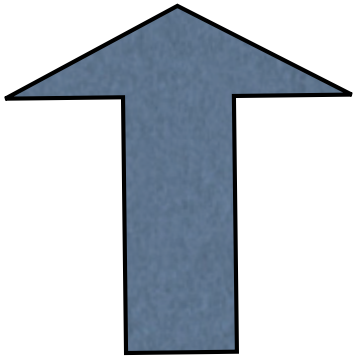
People don't always look where low-level models predict

(a) Original image (b) Itti and Koch Saliency Map (c) eye tracking locations

Where people look



- Top down task and scene dependent cues
- Bottom up saliency cues



Our Learning Approach

- Goal: Learn a model of where people look directly from eye tracking data
- Steps
  - Collect eye tracking dataset
  - Learn a new model

Collecting a database of eye tracking data

Collect eye tracking data

Natural images of objects and scenes downloaded from Flickr and LabelMe

Collect eye tracking data

screen resolution
1280x1024

each image shown for
3 seconds

eye tracker measures
location of eye fixation
several times a second.

user rests head in chin rest

[Photo Credit: Jason Dorfman CSAIL website]

15 users on 1003 images

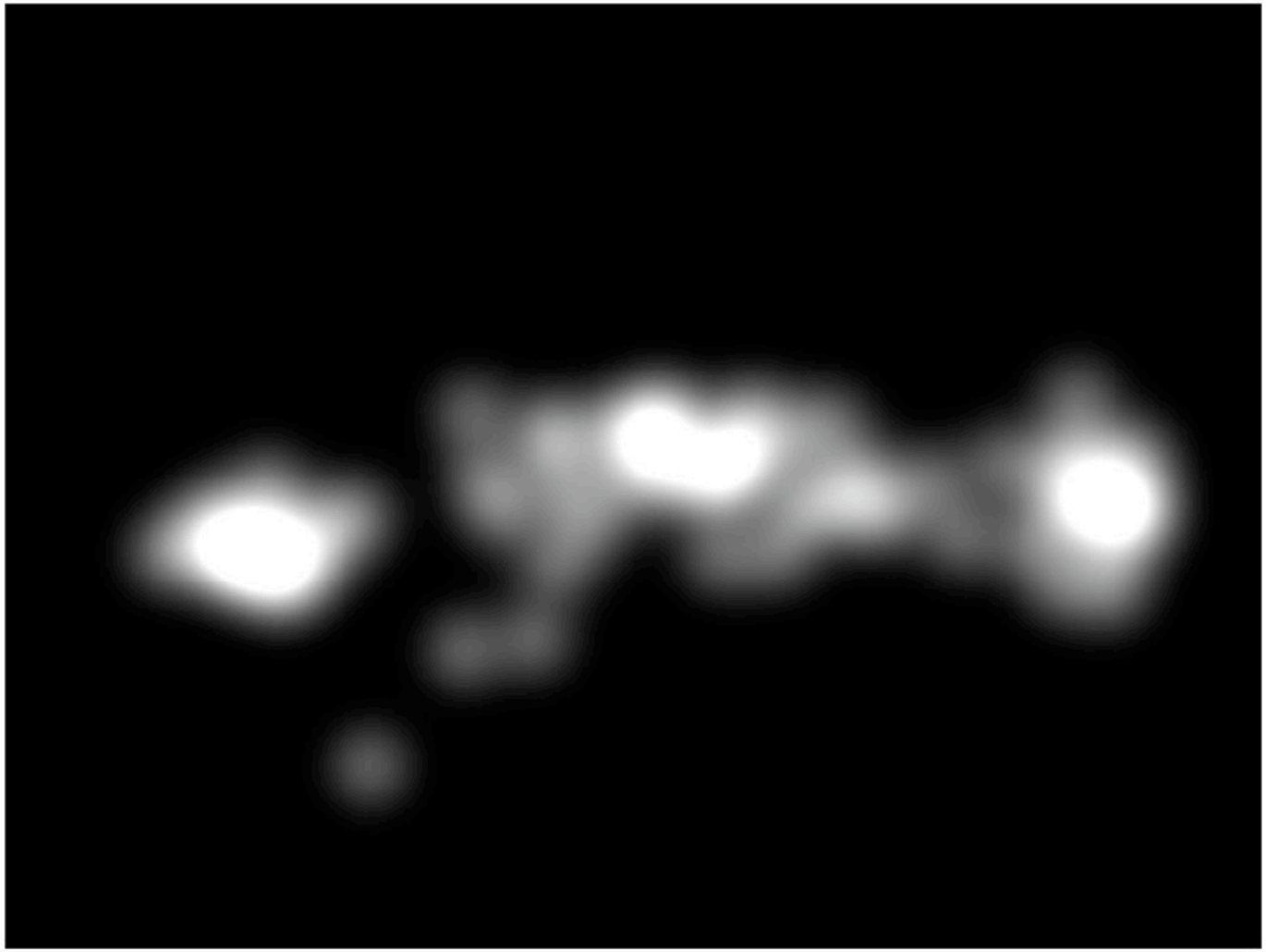
Eye tracking data

fixations for one user

Eye tracking data

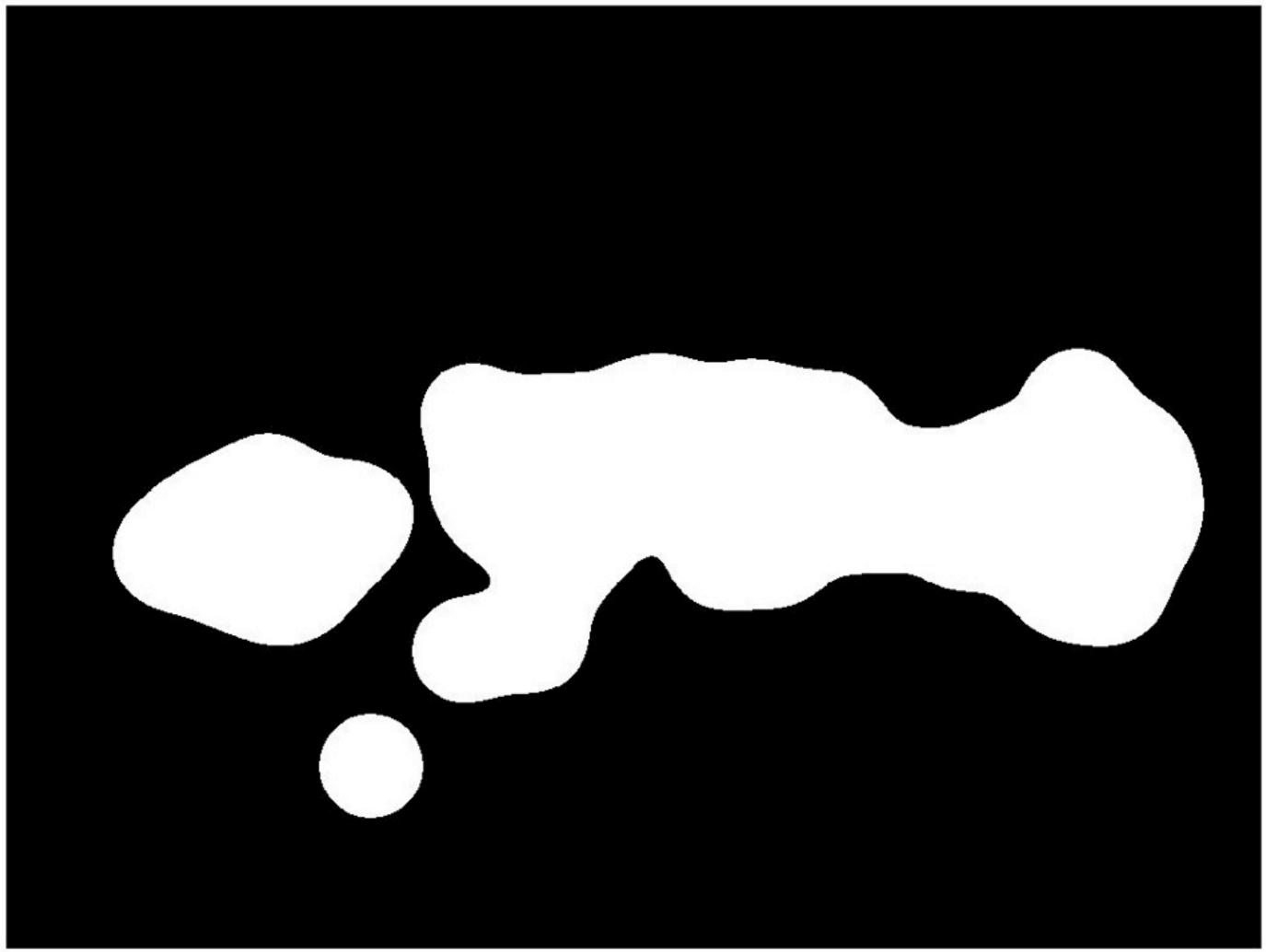
first 5 fixations for 15 users

Eye tracking data



Average fixation locations / continuous saliency map

Eye tracking data

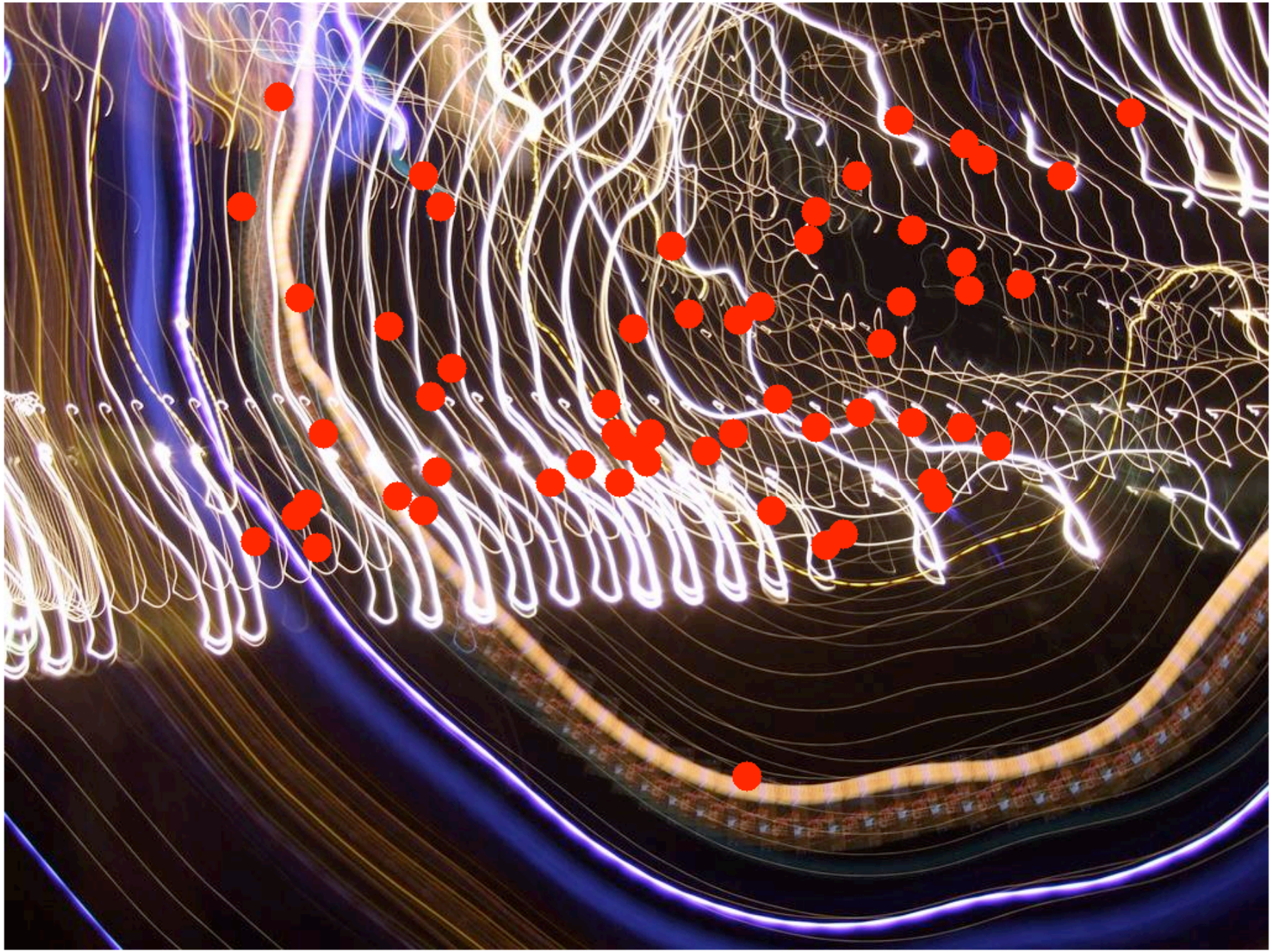


top 20% salient locations

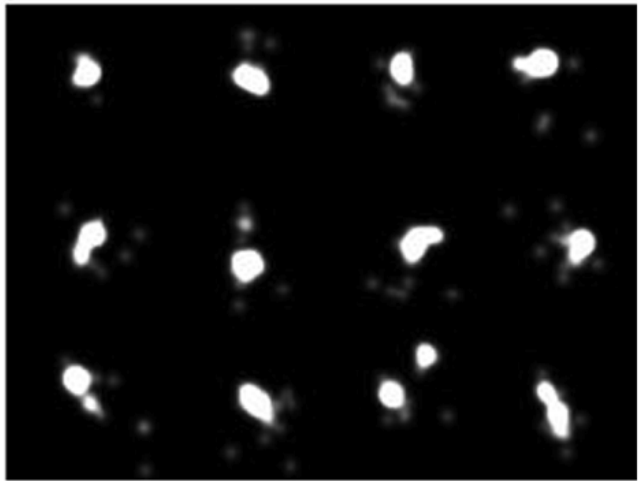
Where do you look?

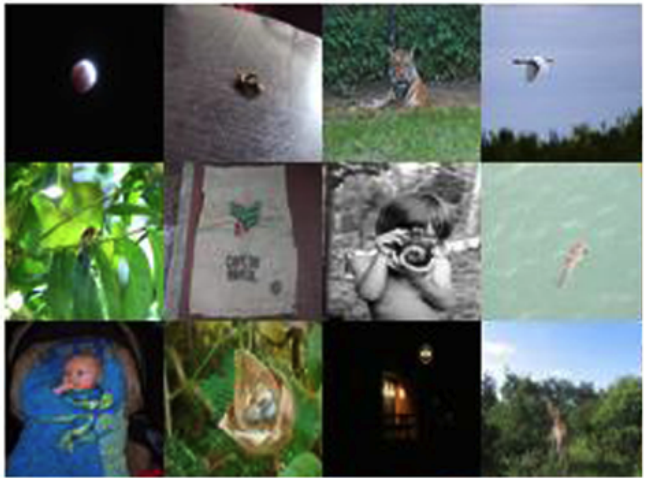

cu motor cu injecție
MARGHITA
360113
AMBULANTA
TRANSPORT
BIHOR-ORADEA
SERVICIUL JUDETEAN DE AMBULANTA


cu motor cu injecție
BIHOR-ORADEA
SERVICIUL JUDETEAN DE AMBULANTA
BH 46D

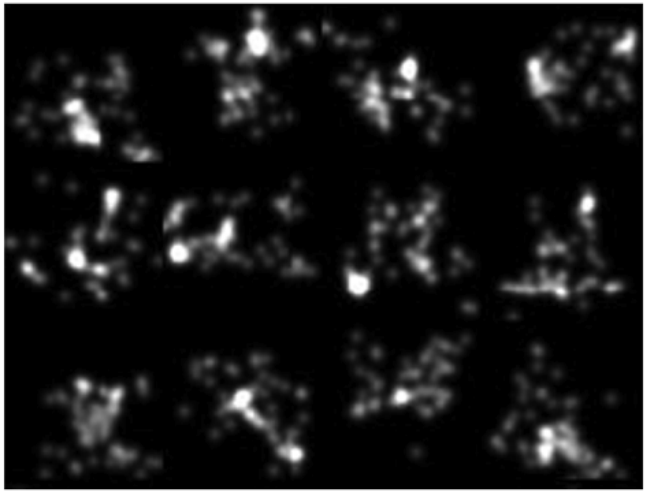


How consistent are humans?



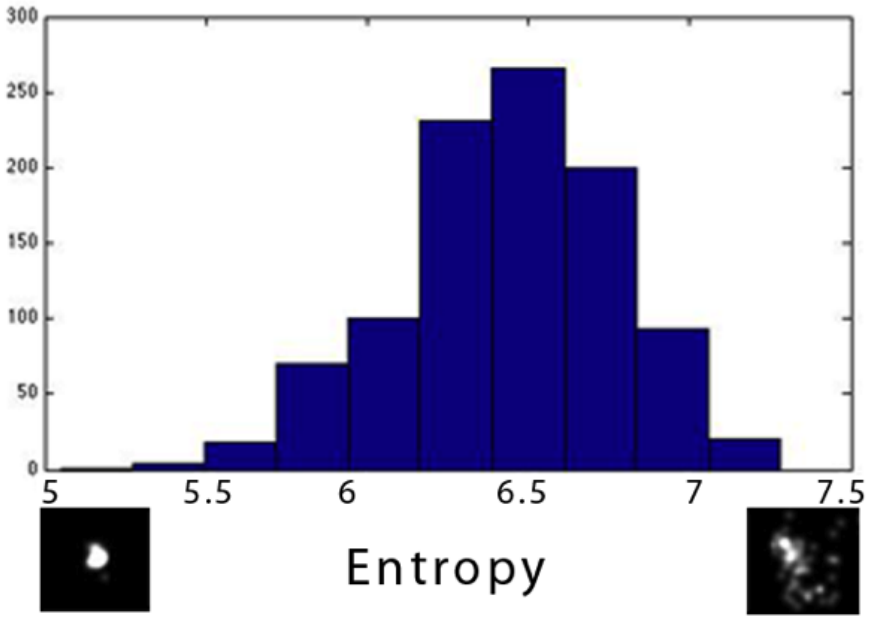


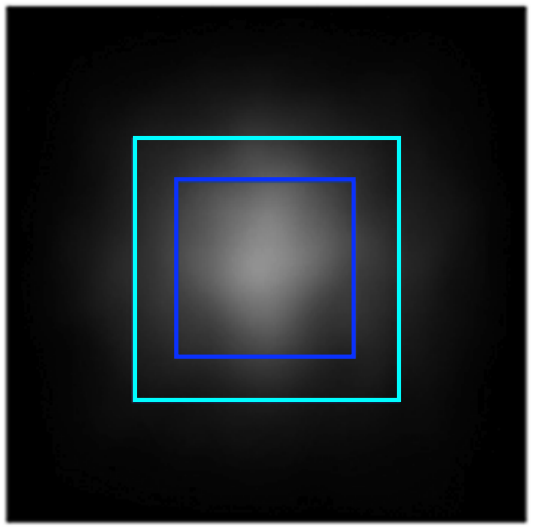
Low entropy saliency maps



High entropy saliency maps

Bias towards the center

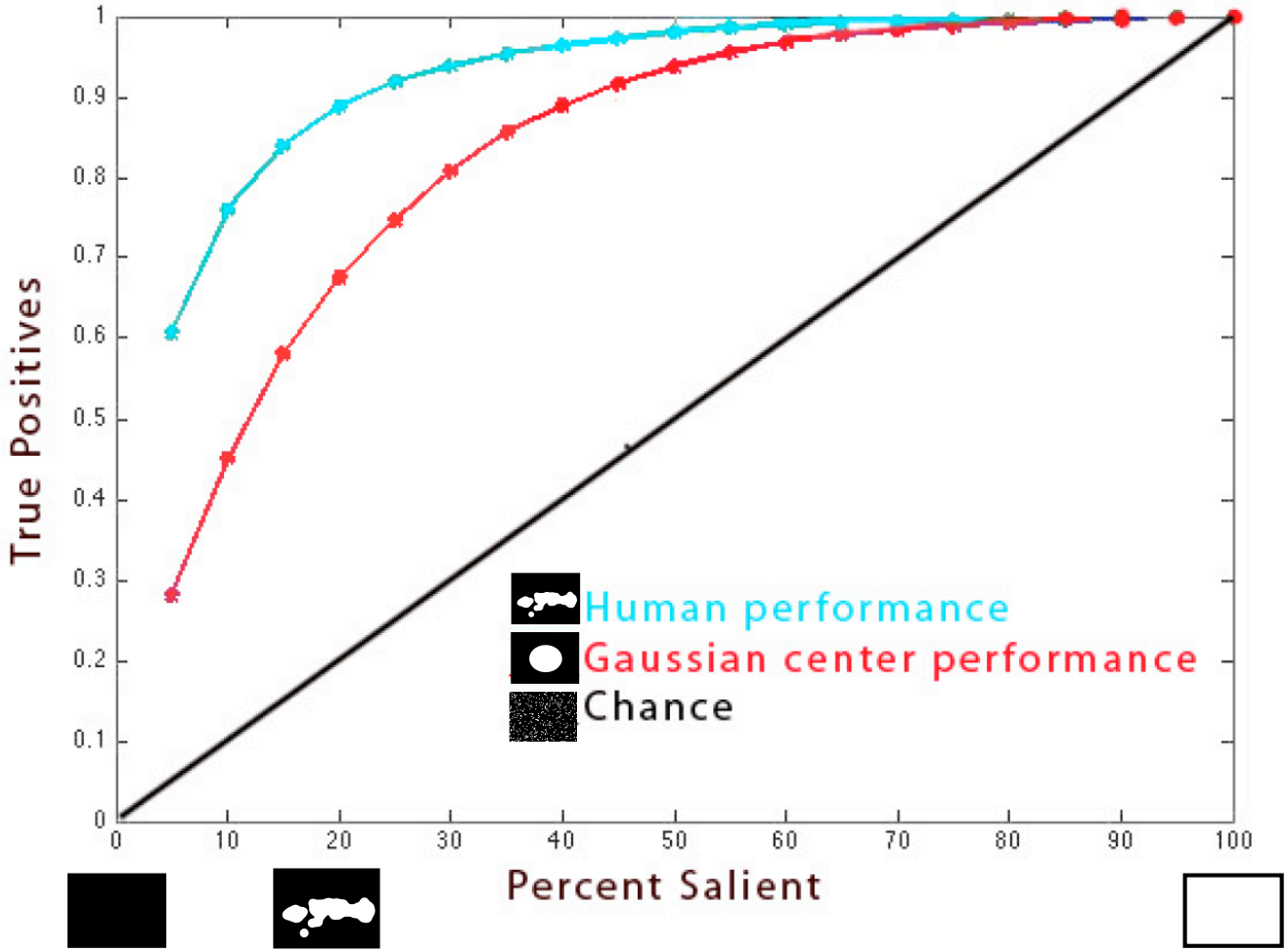


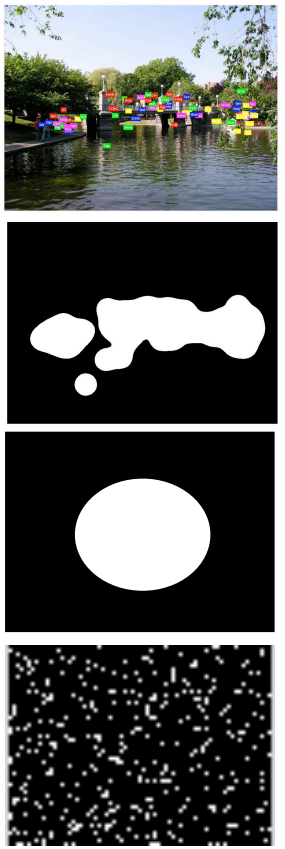


Avg of all saliency maps

40% of fixations within the center 11% of image
70% of fixations within the center 25% of image

Human performance





Where people look

Faces

People

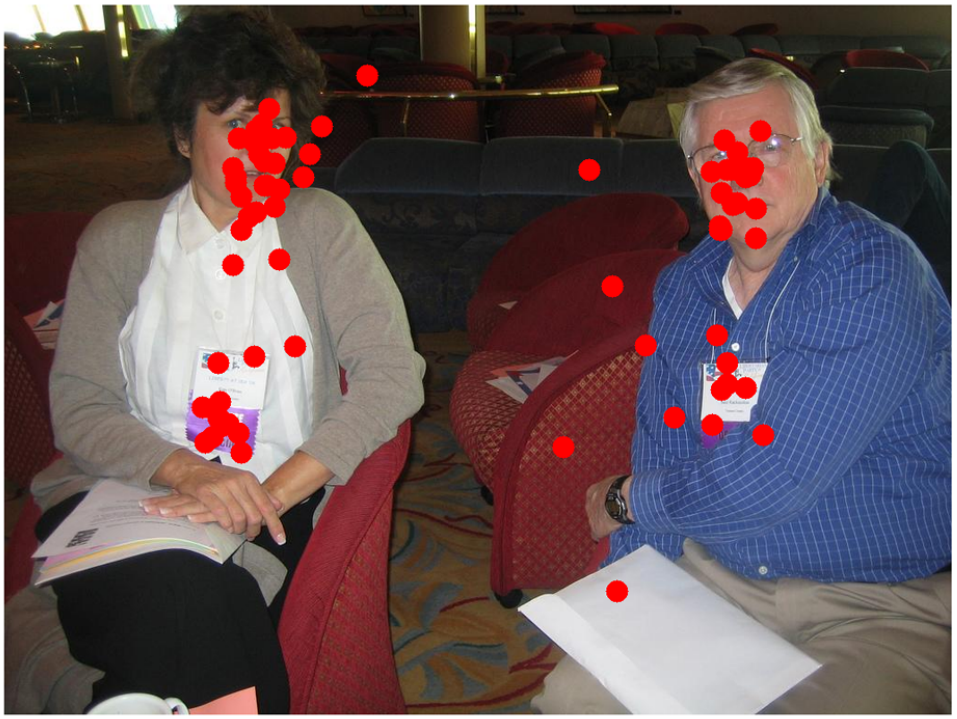
Text

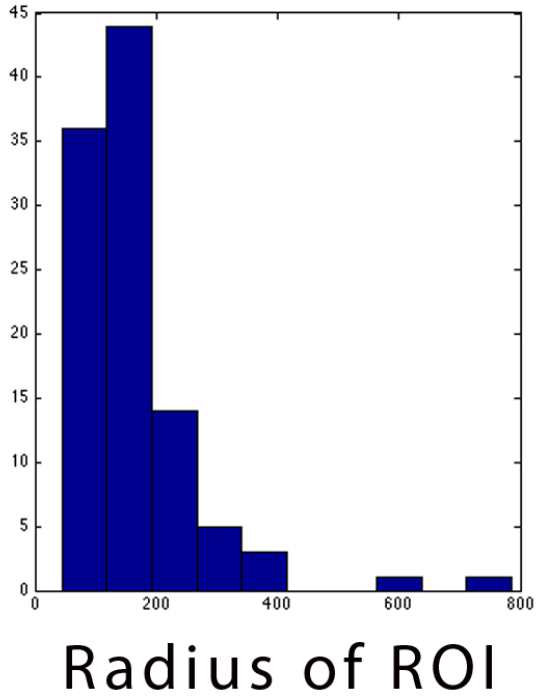
Body parts

Cars

Animals

Size of Region of Interest





Learning a model

Features

- Low level
  illuminance, orientation, color
- Mid level?
  vanishing point, horizon line
- High level
  face detection, object detection

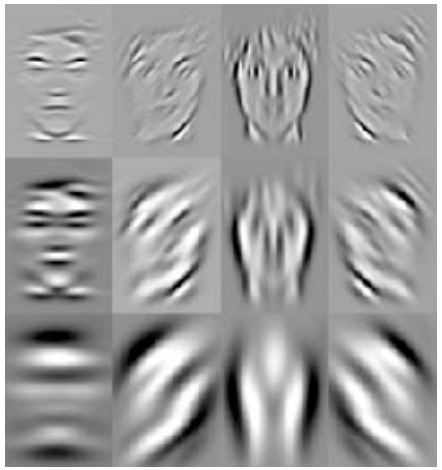
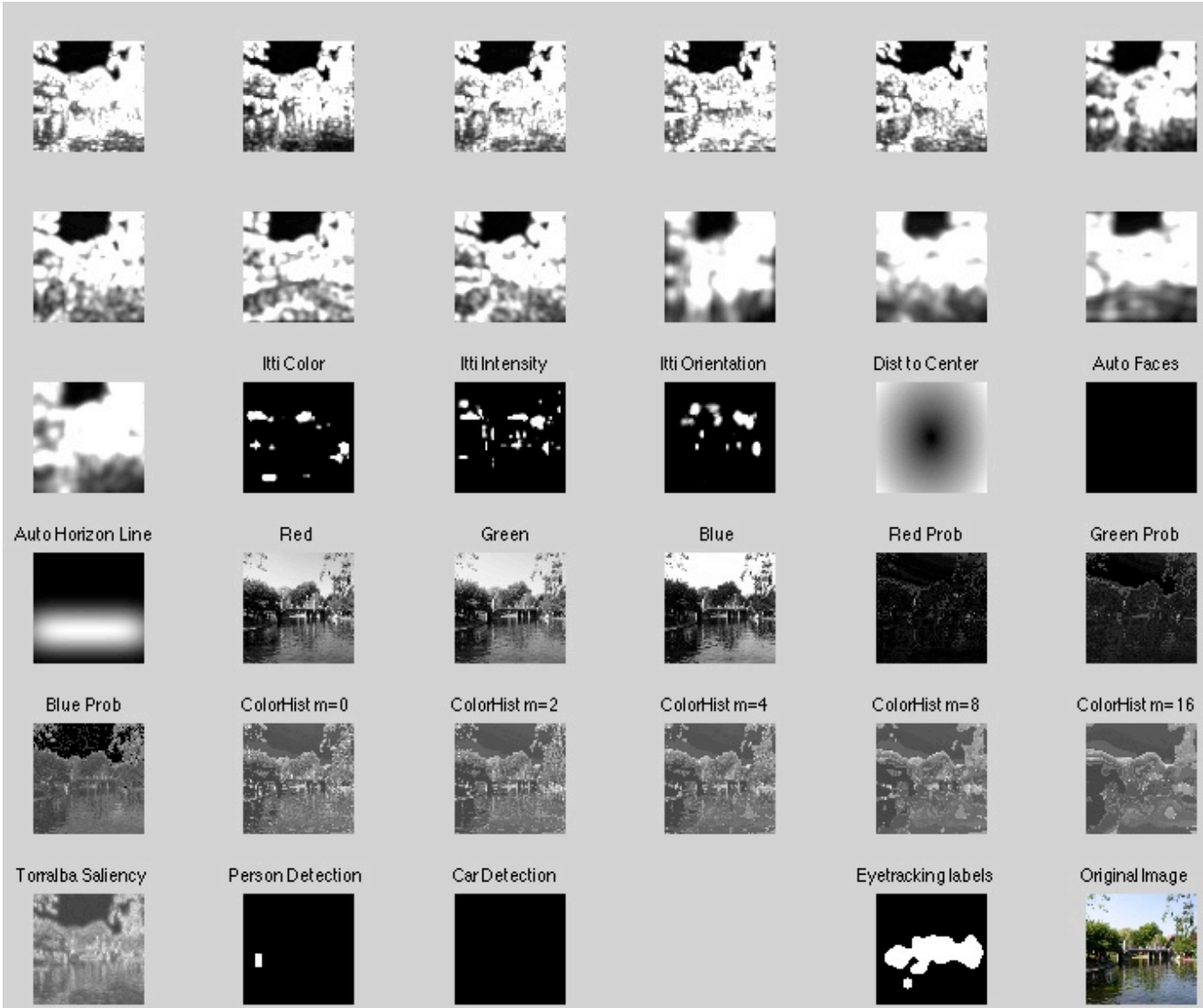


Image filtered with Difference-of-Gaussian(DoG) filters

Viola Jones Face detector

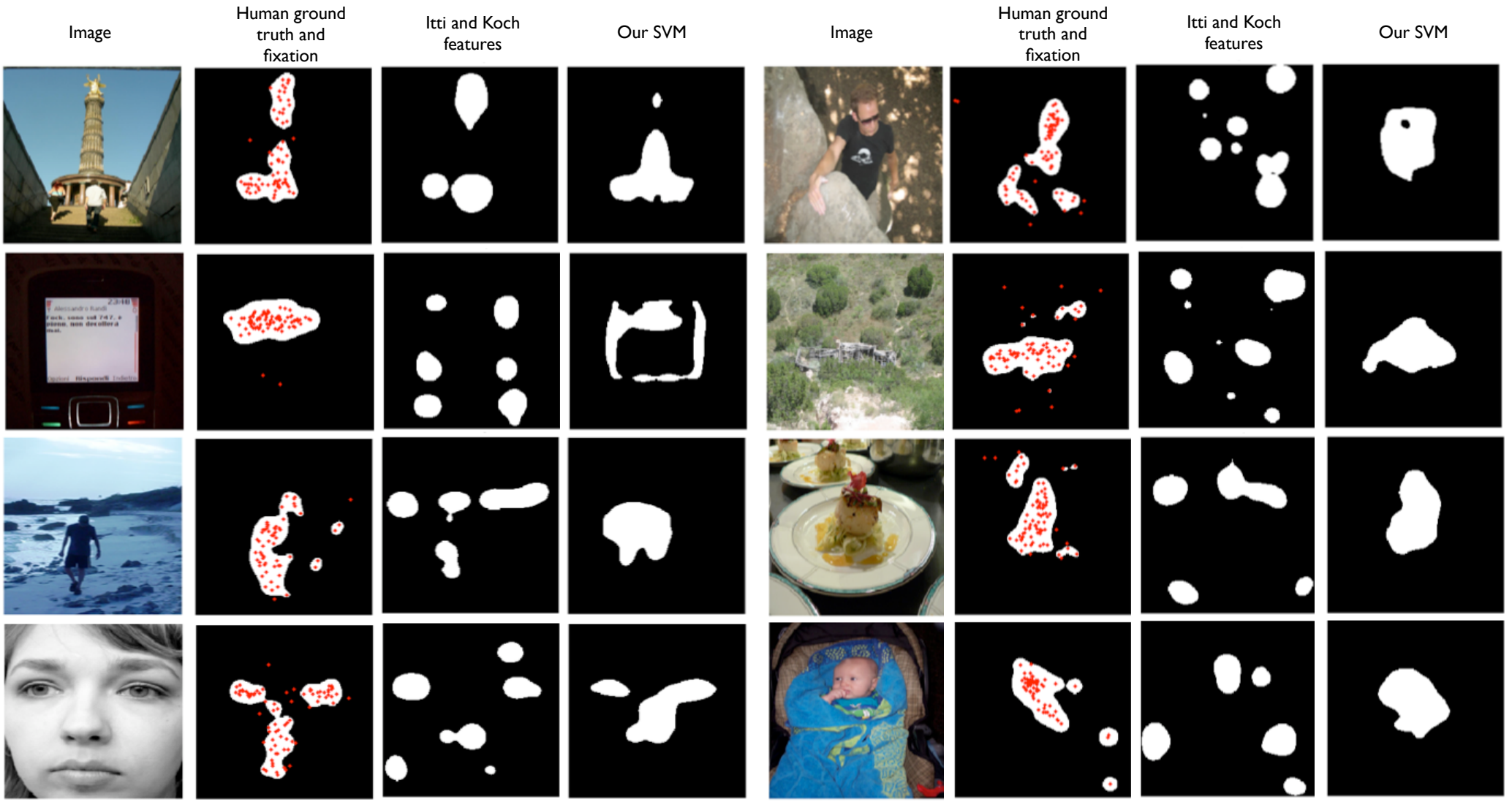
Features

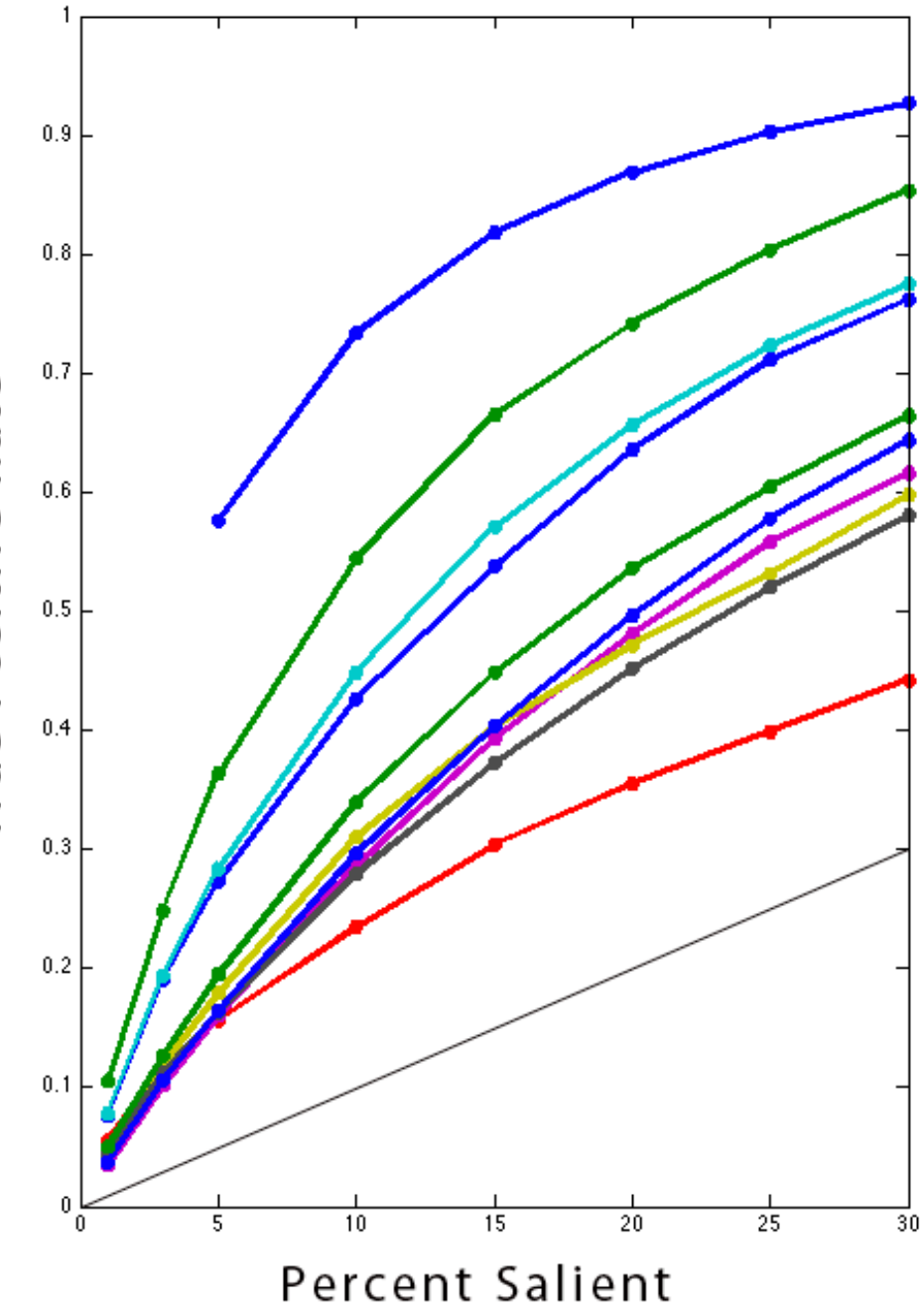

Itti Color
Itti Intensity
Itti Orientation
Dist to Center
Auto Faces
Auto Horizon Line
Red
Green
Blue
Red Prob
Green Prob
Blue Prob
ColorHist m=0
ColorHist m=2
ColorHist m=4
ColorHist m=8
ColorHist m=16
Torralba Saliency
Person Detection
Car Detection
Eyetracking labels
Original Image

Learn a model

- collect positive salient and negative non salient examples (10 pos : 10 neg)
- Linear support vector machine
- Test on single features and all features

Saliency maps from different models





Humans

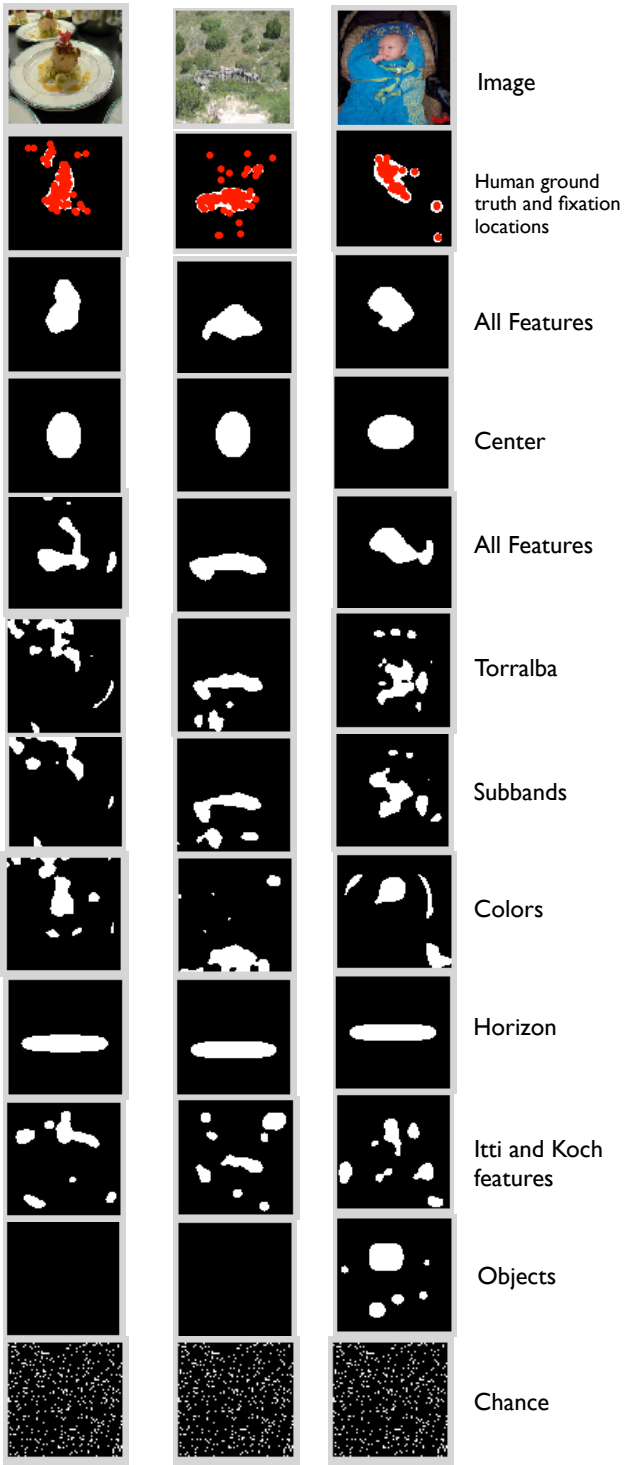
All Features (our model)

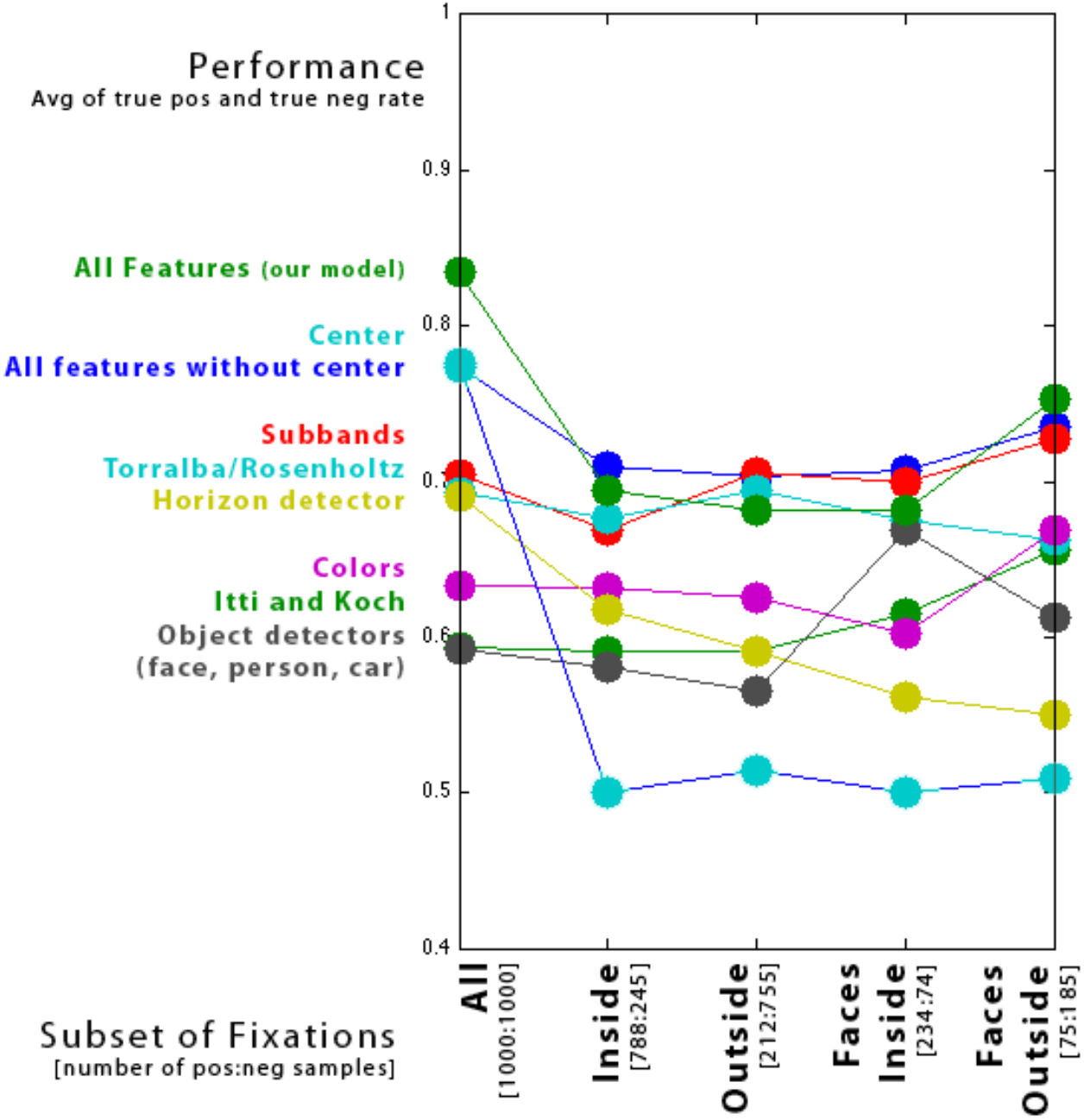
Center
All features without center

Torralba/Rosenholtz
Subbands
Colors
Horizon detector
Itti and Koch

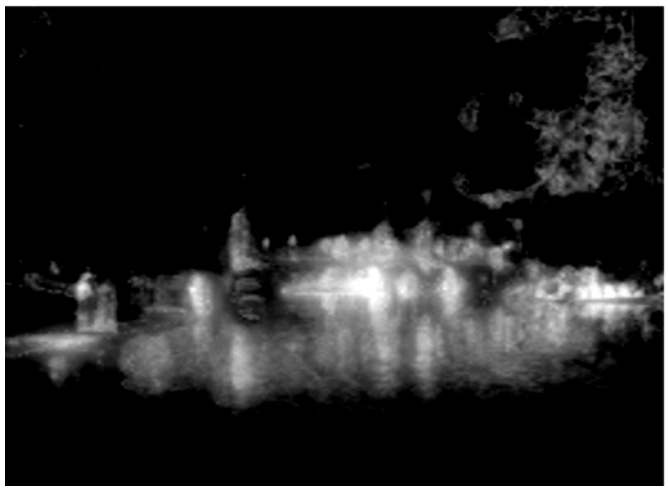
Object detectors (face, person, car)

Chance




Performance
Avg of true pos and true neg rate
All Features (our model)
Center
All features without center
Subbands
Torralba/Rosenholtz
Horizon detector
Colors
Itti and Koch
Object detectors
(face, person, car)
1
0.9
0.8
0.7
0.6
0.5
0.4
All [1000:1000]
Inside [788:245]
Outside [212:755]
Faces Inside [234:74]
Faces Outside [75:185]
Subset of Fixations
[number of pos:neg samples]

NPR application



Conclusion

- Have new large collection of eye tracking data
- Can learn models of saliency and BEYOND
- Future work: enhance model, applications for saliency, explore cropping

Questions?