Learning to Detect Natural Image Boundaries Using Local Brightness, Color and Texture Cues

David Martin, Charless Fowlkes, Jitendra Malik
Computer Science Division
University of California at Berkeley

Presented by Jean-Francois Lalonde

Disclaimer: Most of the slides are from D. Martin

What is a Boundary?

Original

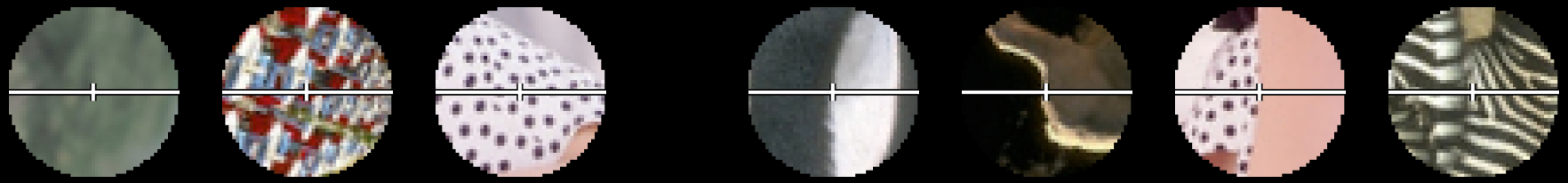
Canny

Human

http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/

How do humans do this?

- Low-level cues?
  - Brightness? Color? Texture?
- Mid-level cues?
  - Continuity? Closure? Symmetry?
- High-level cues?
  - Context? Object recognition?
- This paper: what is the optimal way to use **LOCAL** information?



- Psychophysics of localization:
  - Multi-Attribute Boundaries [Rivest/Cavanagh 1996]
    - luminance, color, motion, texture
    - Information pooled prior to localization
  - Texture Boundaries [Landy/Kojima 2001]
    - frequency, orientation, contrast
- Their approach: Supervised learning to optimally combine boundary cues.

Dataflow

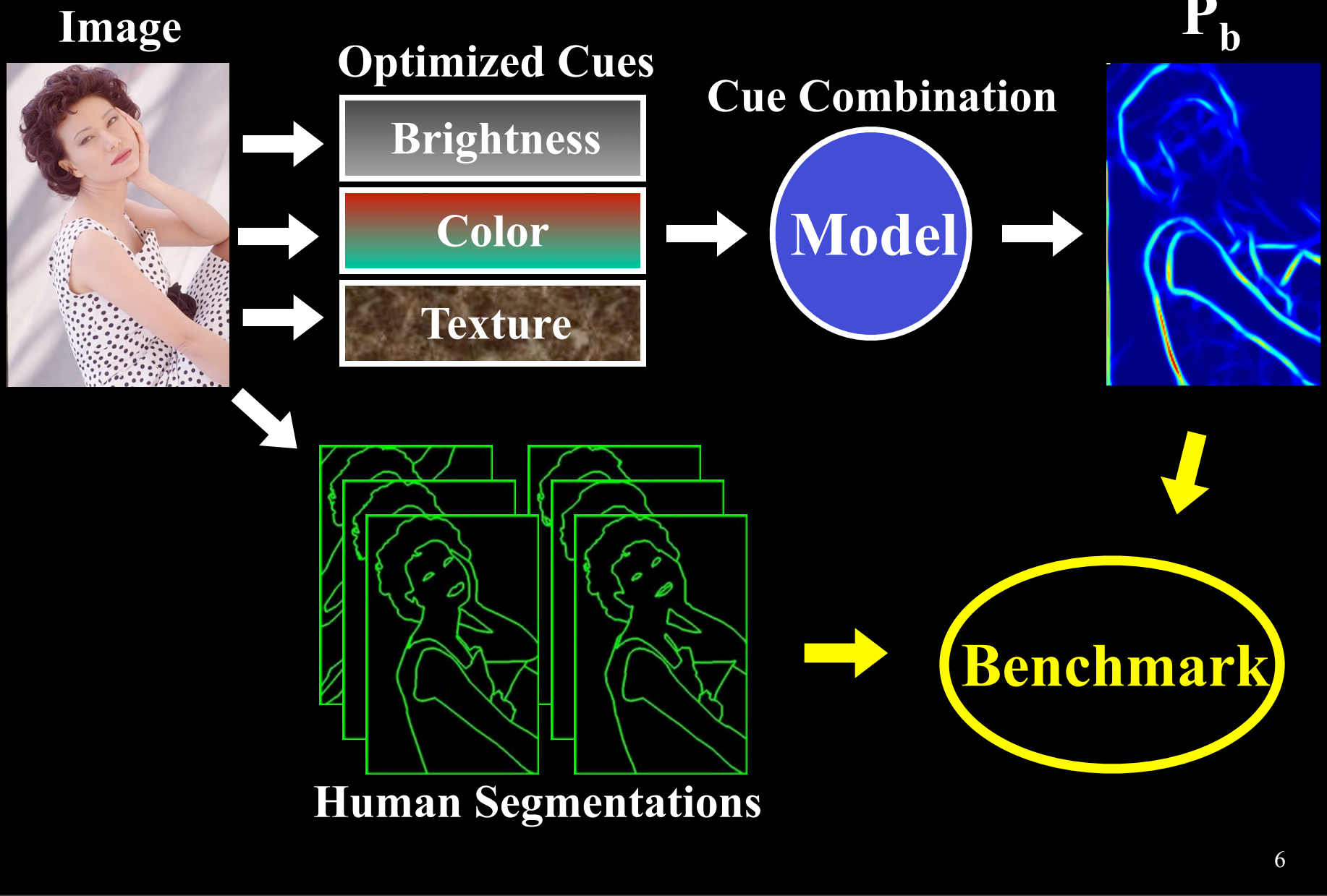
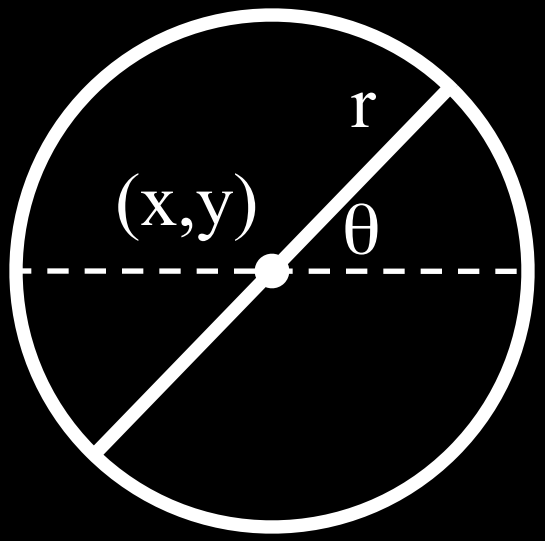


Image
Optimized Cues
Brightness
Color
Texture
Cue Combination
Model
P_b
Human Segmentations
Benchmark

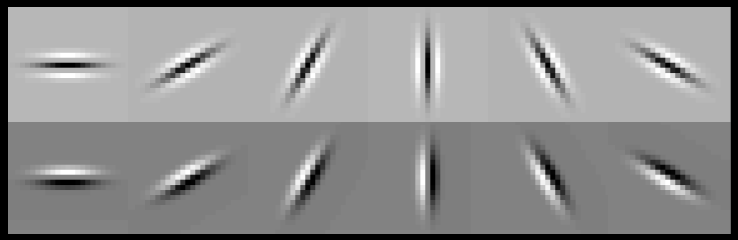
Brightness and Color Features

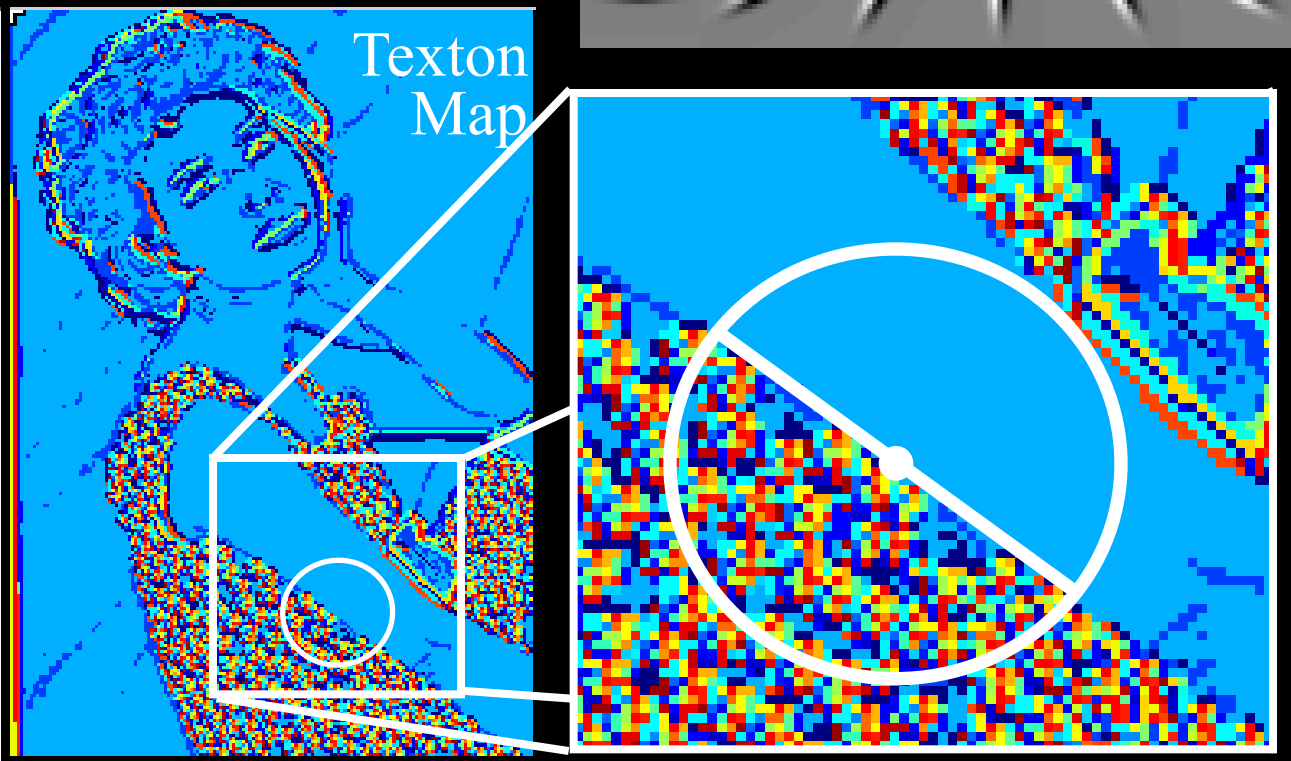
- 1976 CIE L*a*b* color space
- Brightness Gradient BG(x,y,r,θ)
  - χ^2 difference in L* distribution
- Color Gradient CG(x,y,r,θ)
  - χ^2 difference in a* and b* distributions

$$\chi^2(g,h) = \frac{1}{2}\sum_i \frac{(g_i - h_i)^2}{g_i + h_i}$$

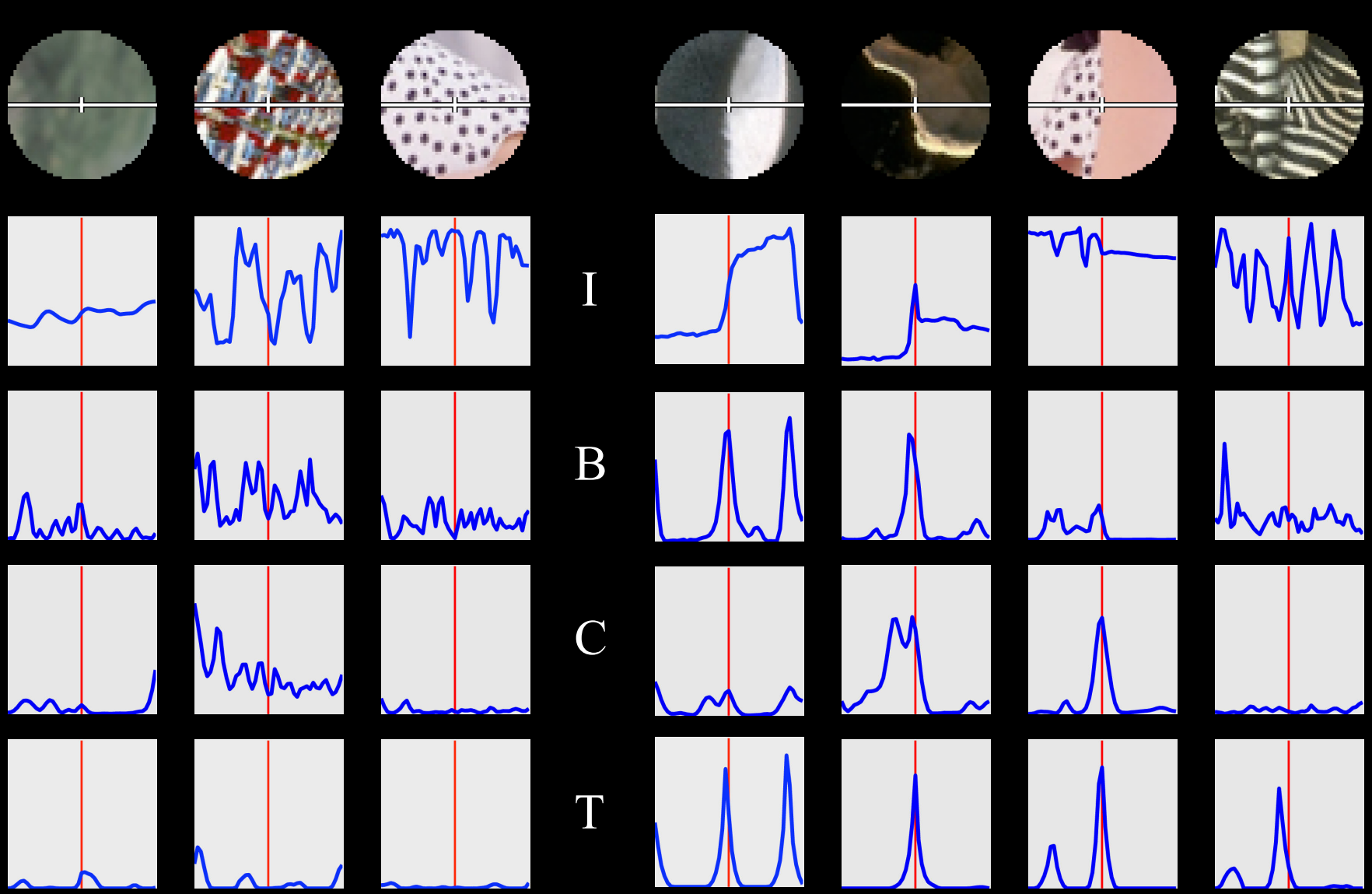


Texture





- Texture Gradient TG(x,y,r,θ)
  - χ^2 difference of texton histograms
  - Textons are vector-quantized filter outputs

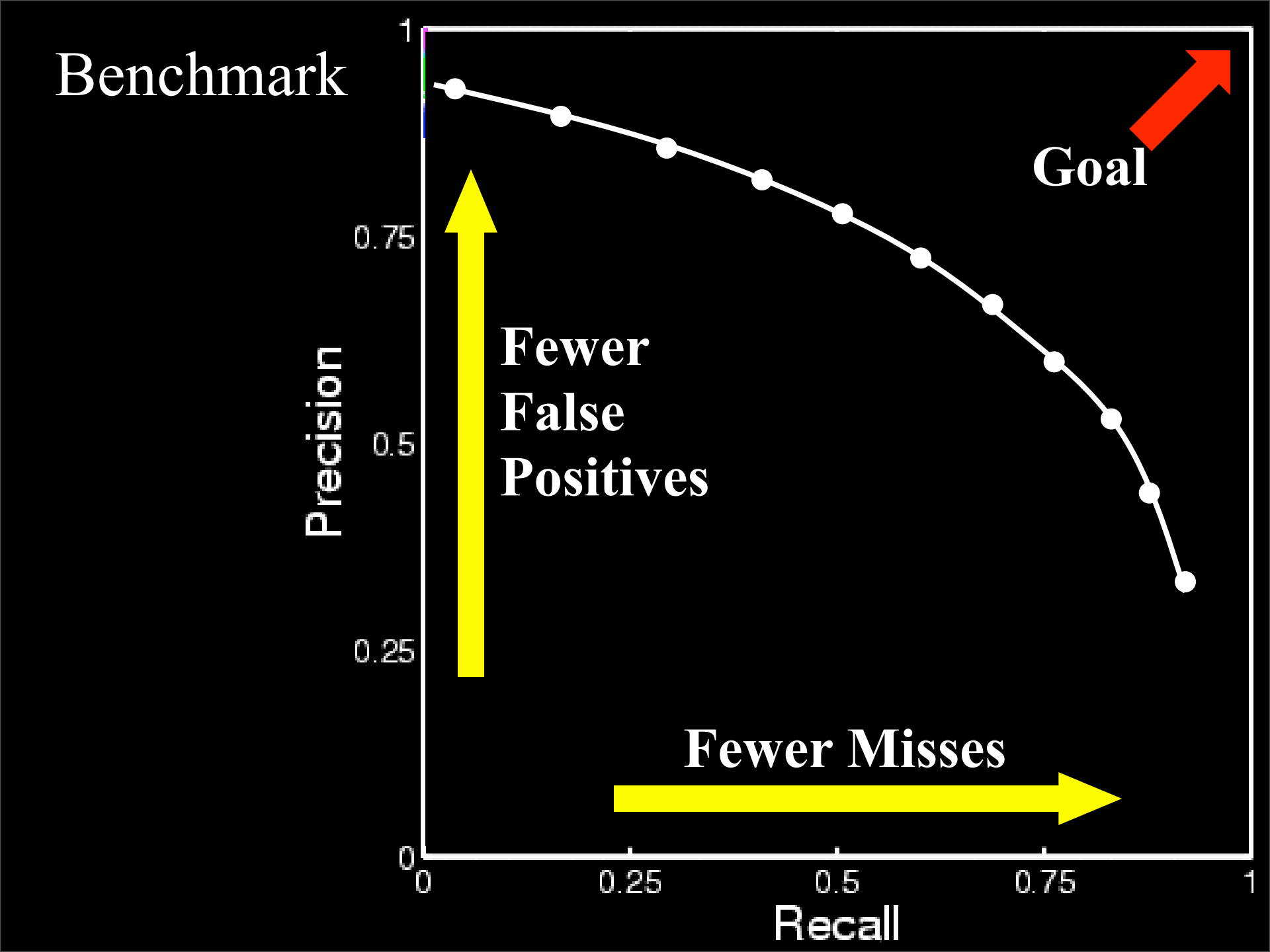

Non-Boundaries
Boundaries
I
B
C
T

Cue Combination Models

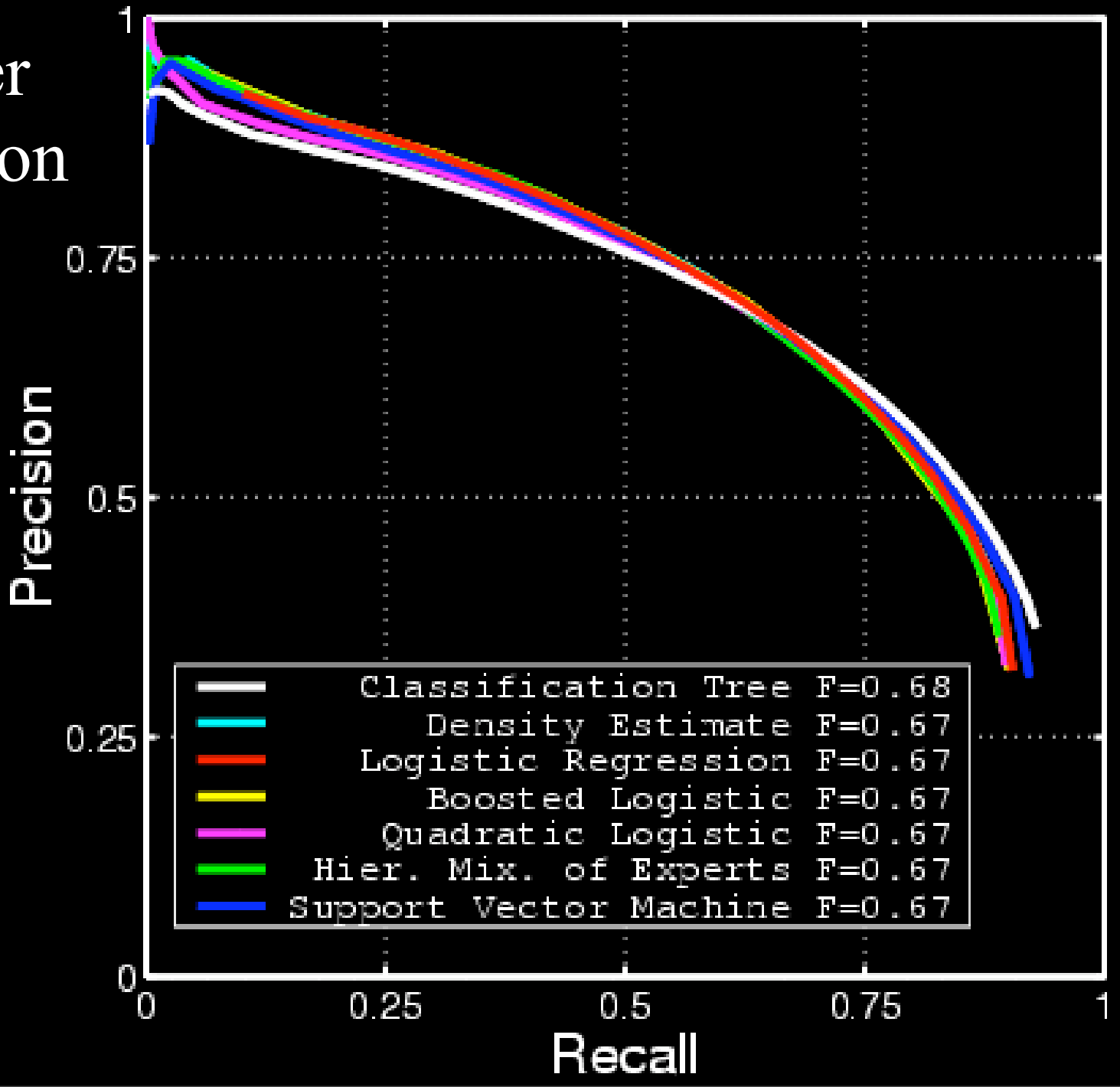
- Classification Trees
  - Top-down splits to maximize entropy, error bounded
- Density Estimation
  - Adaptive bins using k-means
- Logistic Regression, 3 variants
  - Linear and quadratic terms
  - Confidence-rated generalization of AdaBoost (Schapire&Singer)
- Hierarchical Mixtures of Experts (Jordan&Jacobs)
  - Up to 8 experts, initialized top-down, fit with EM
- Support Vector Machines (`libsvm`, Chang&Lin)
  - Gaussian kernel, ν-parameterization

➢ Range over bias, complexity, parametric/non-parametric

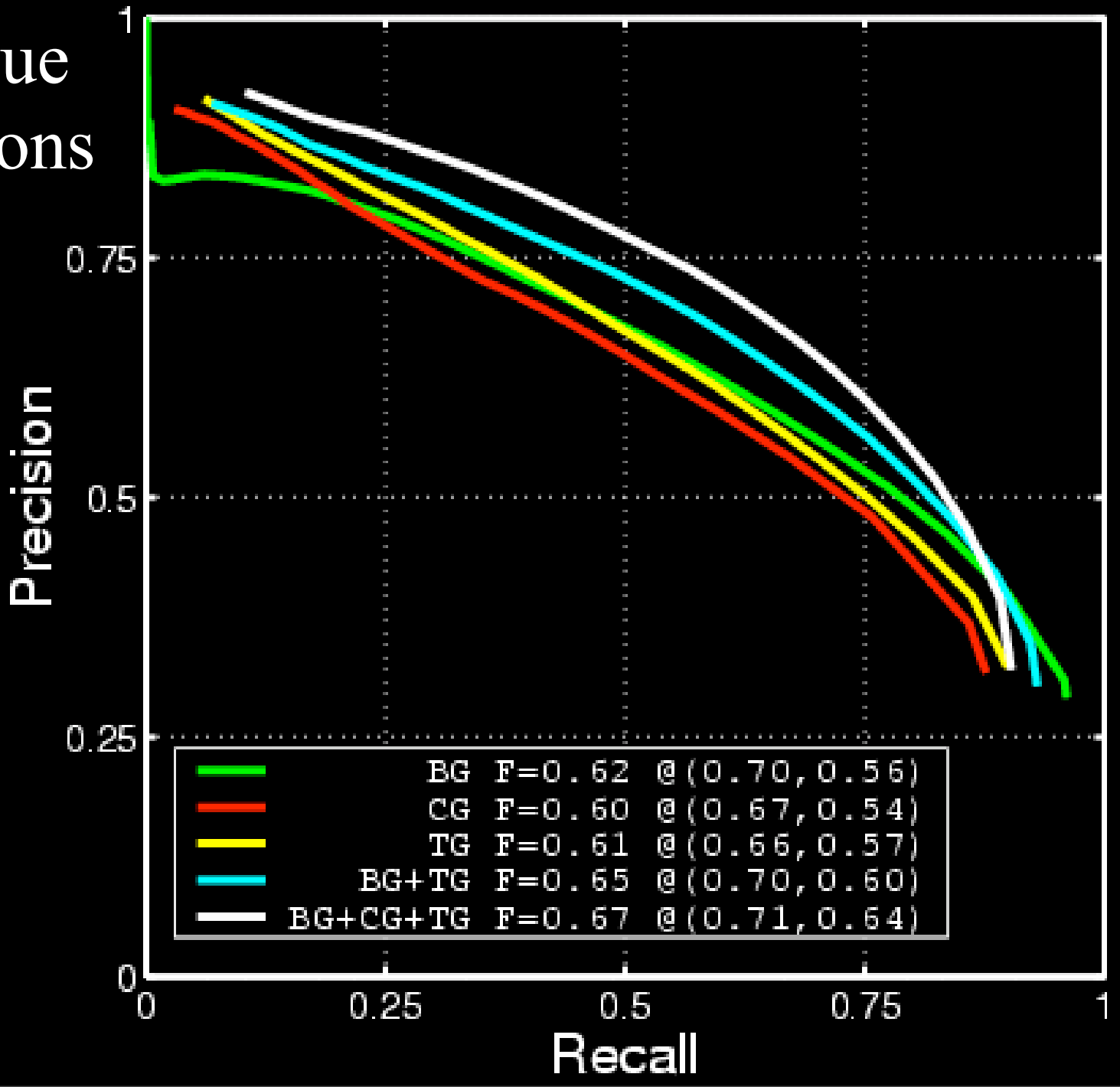
Benchmark


1
0.75
0.5
0.25
0
Precision
0
0.25
0.5
0.75
1
Recall
Goal
Fewer
False
Positives
Fewer Misses

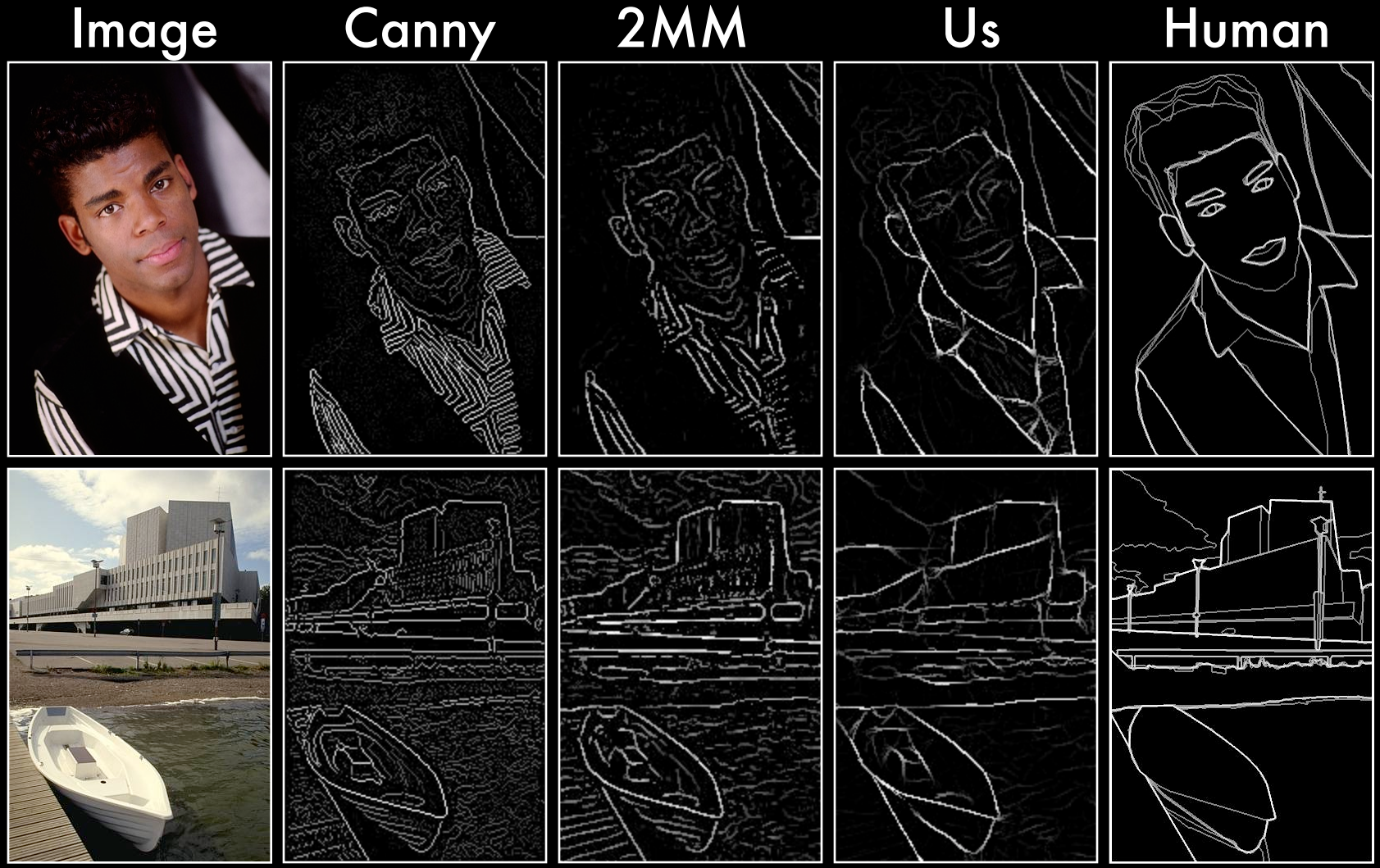
Classifier Comparison



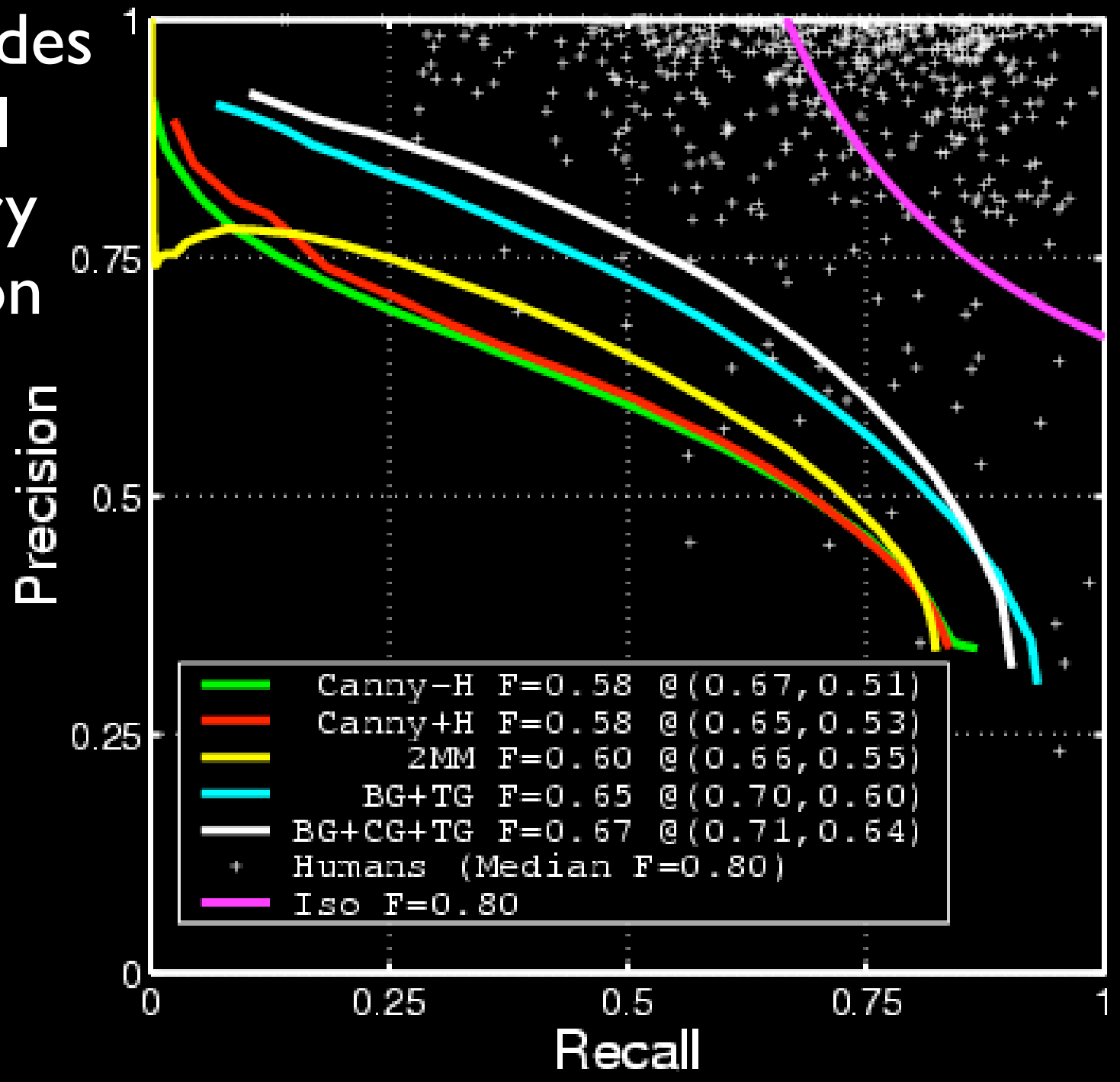
Various Cue Combinations



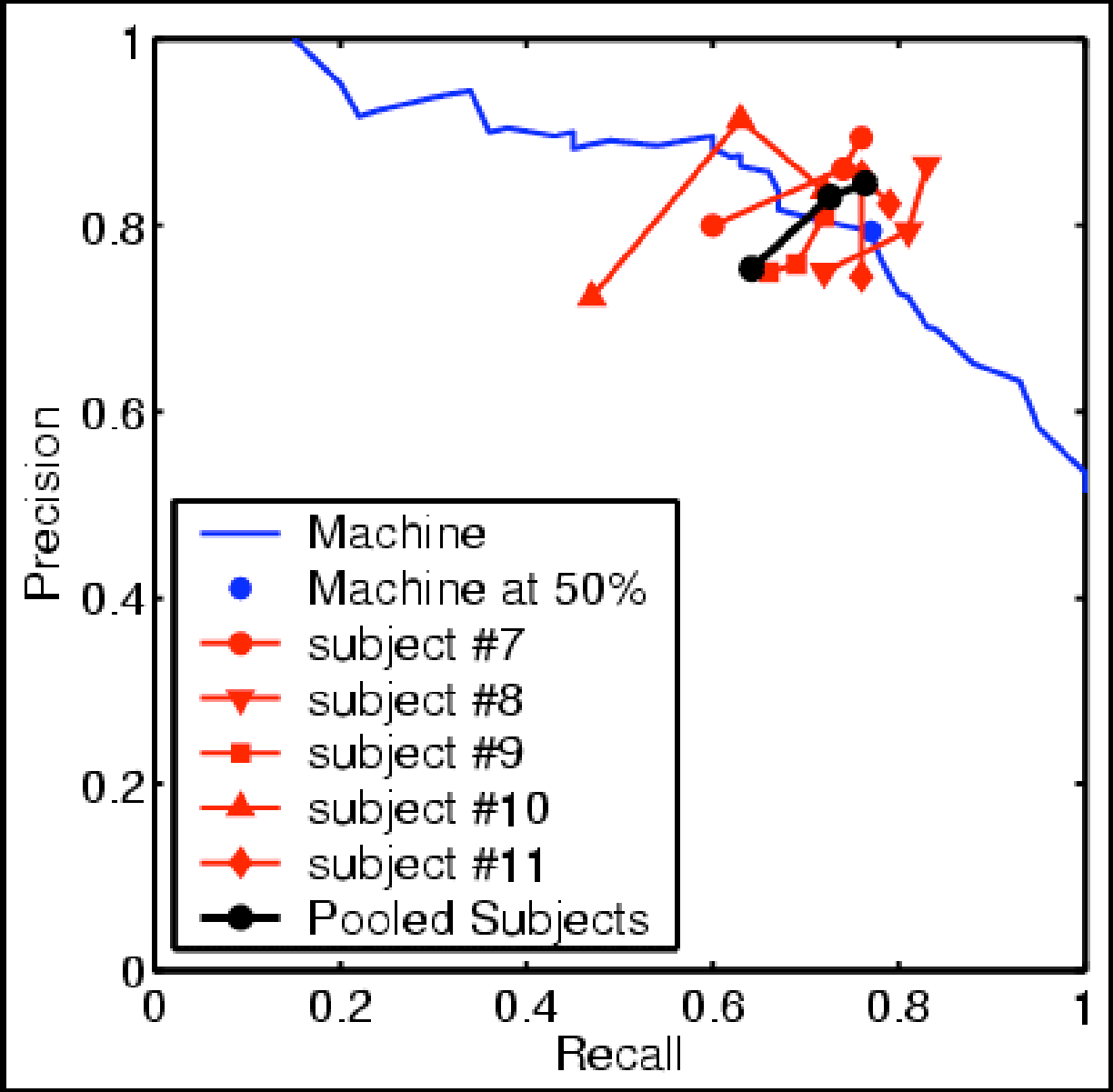
Detector Comparison



Two Decades of Local Boundary Detection



Human vs machine on local patches



Conclusion

1. A simple linear model is sufficient for cue combination
   - All cues weighted approximately equally in logistic
   - Linear model supported by psychophysics
2. Texture gradients are a powerful and necessary cue
3. Significant improvement over state-of-the-art in local boundary detection
   - $P_b(x,y,\theta)$ good for higher-level processing
4. Human performance on patches??
   - ECVP'03

The End