The Ecological Statistics of Grouping by Similarity

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

Brunswik & Kamiya 1953

- Thesis: Gestalt rules reflect the structure of the natural world
- Attempted to validate the grouping rule of *proximity of similars*

- Brunswik was ahead of his time… we now have the tools.

Egon Brunswik (1903-1955)

Natural Image Statistics

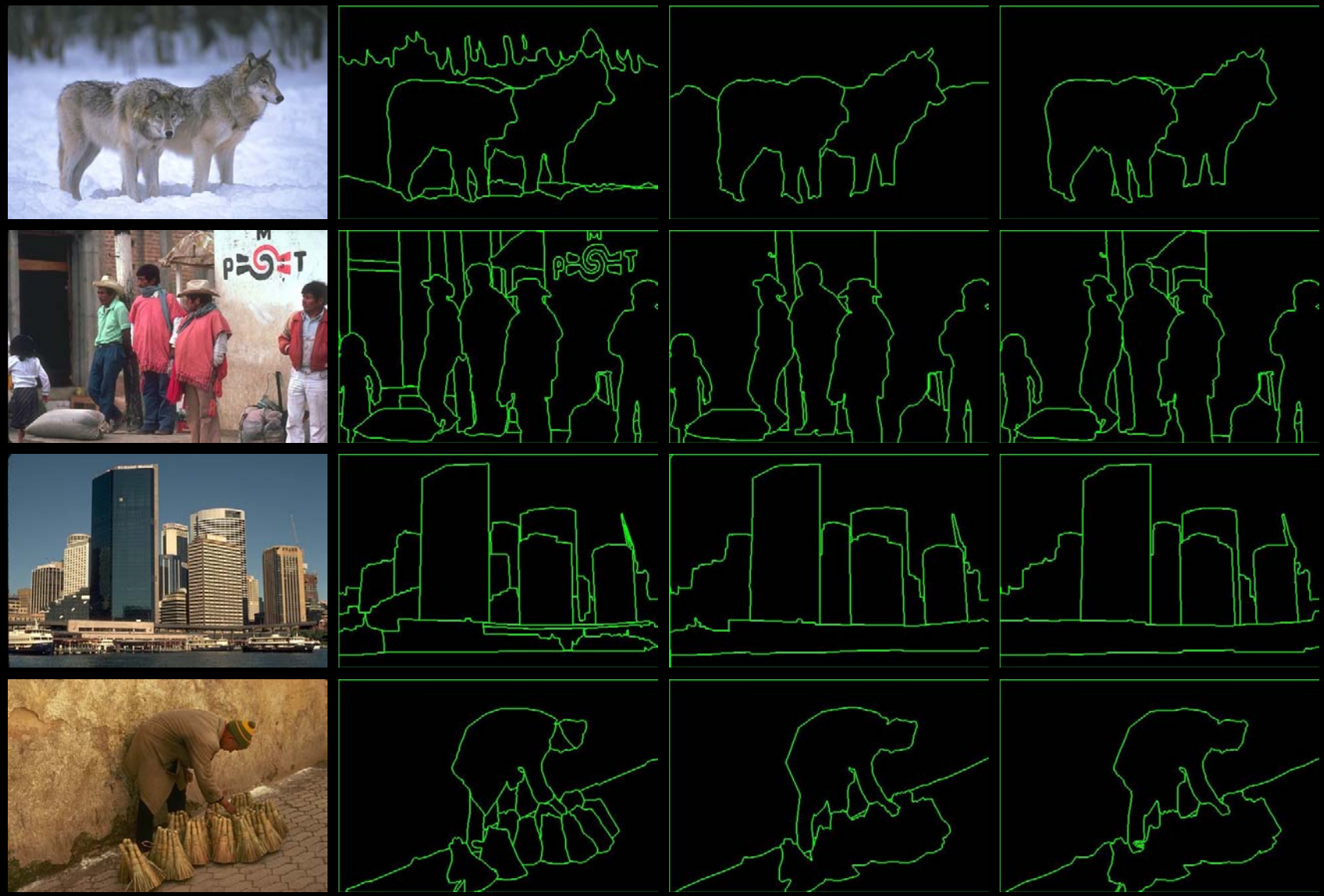
P(image features)

- *Edges/Filters/Coding*: Ruderman 94/97, Dong/Atick 95, Olshausen/Field 96, Bell/Sejnowski 97, Hateren/Schaaf 98, Buccigrossi/Simoncelli 99, Alvarez/Gousseau/Morel 99, Huang/Mumford 99

P(grouping | image features)

- *Proximity*: Brunswik/Kamiya 53
- *Good Continuation*: Geisler, et. al. 01, Ren/Malik 02
- *Similarity*: Martin/Fowlkes/Malik 01, and this talk

Human Segmentation Dataset



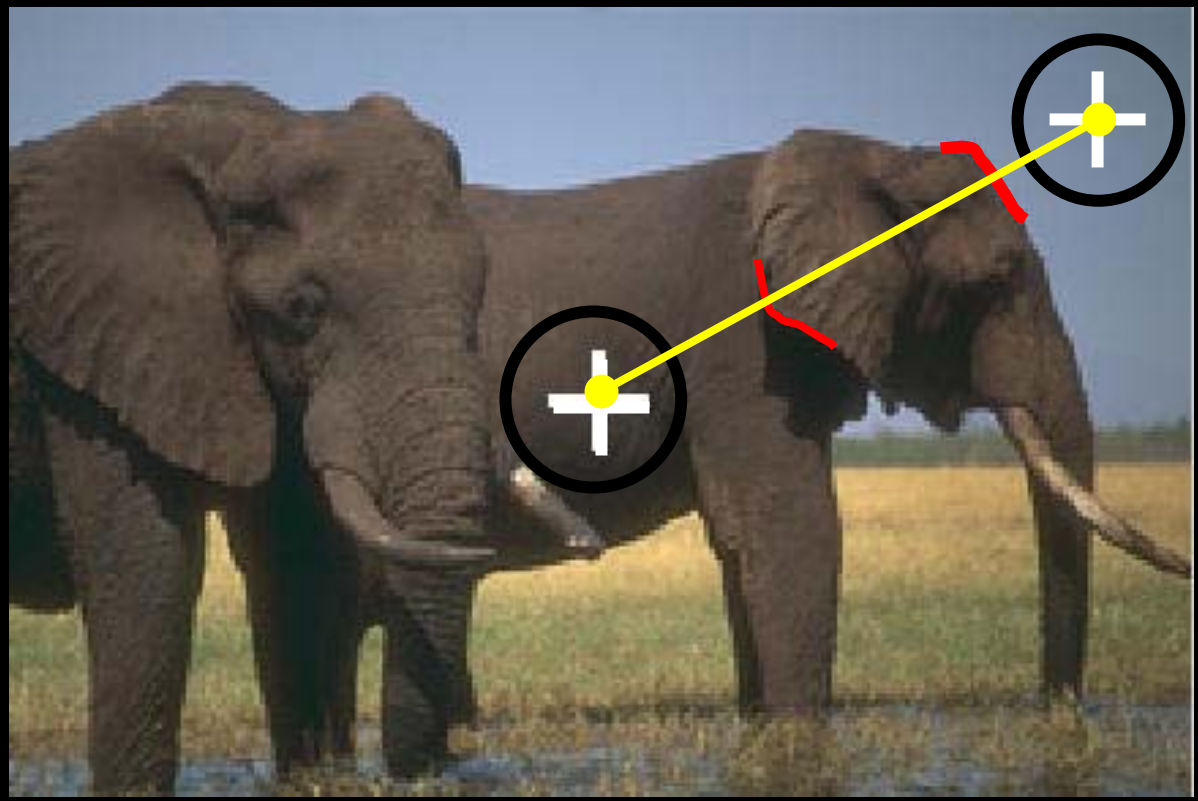
Details…

"You will be presented a photographic image. Divide the image into some number of segments, where the segments represent "things" or "parts of things" in the scene. The number of segments is up to you, as it depends on the image. Something between 2 and 30 is likely to be appropriate. It is important that all of the segments have approximately equal importance."

- 30 subjects, age 19-23
- 8 months
- 1,458 person hours
- 1,020 Corel images
- 11,595 Segmentations
  - 5,555 color, 5,554 gray, 486 inverted/negated

Similarity Cues

a) distance
b) patch-based similarity cues
c) edge-based cues



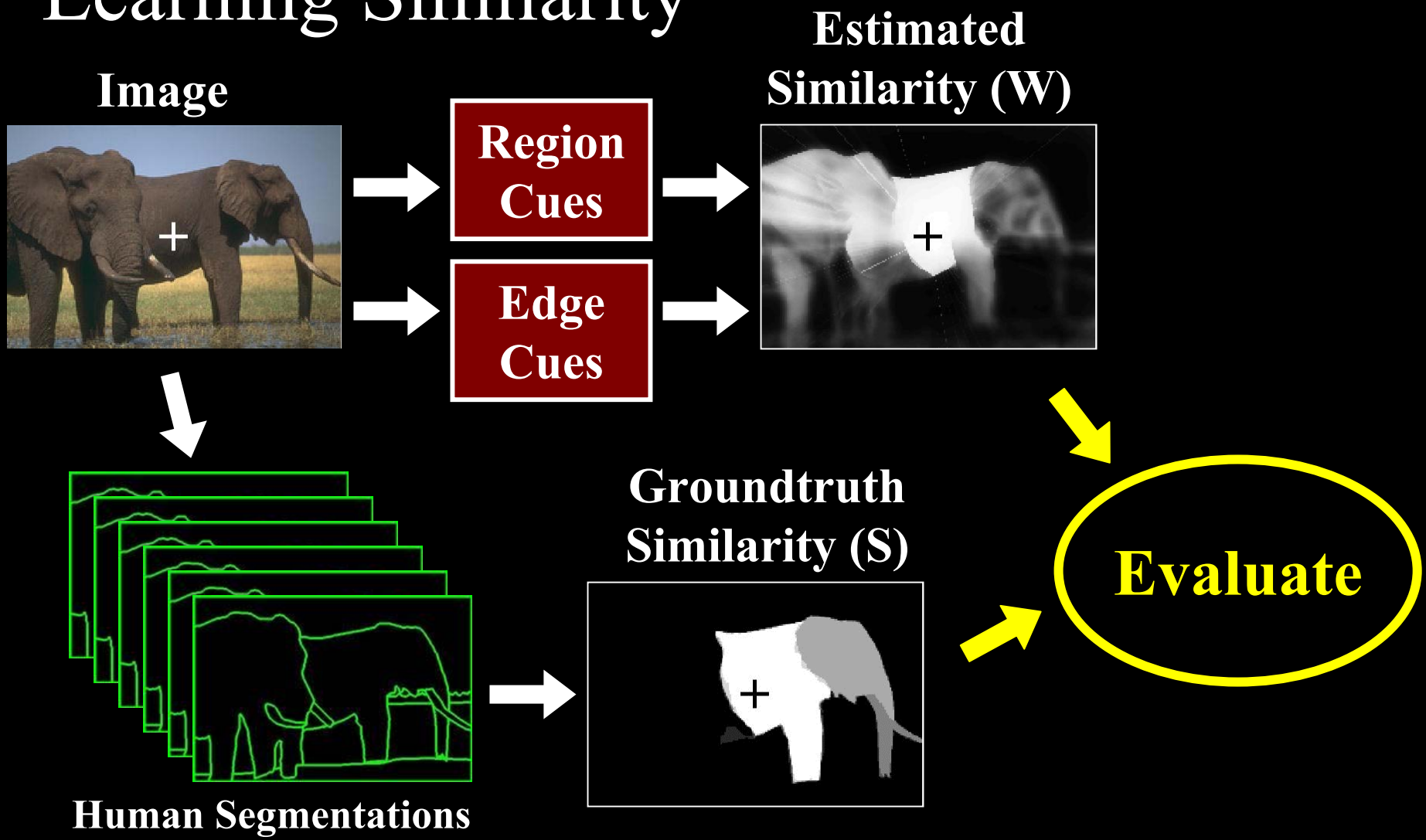
What image measurements allow us to gauge the probability that pixels i and j belong to the same segment?

Brightness, Color and Texture Features

- Brightness and Color Features
  - CIE L*a*b* color-space
  - Estimate distributions of L*, a* and b* values inside analysis window
- Texture Features
  - Filter image with even/odd-symmetric filters which resemble V1 receptive fields.
  - Estimate distribution of vector-quantized filter responses inside analysis window
- Compare histograms using chi-squared difference

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

Learning Similarity



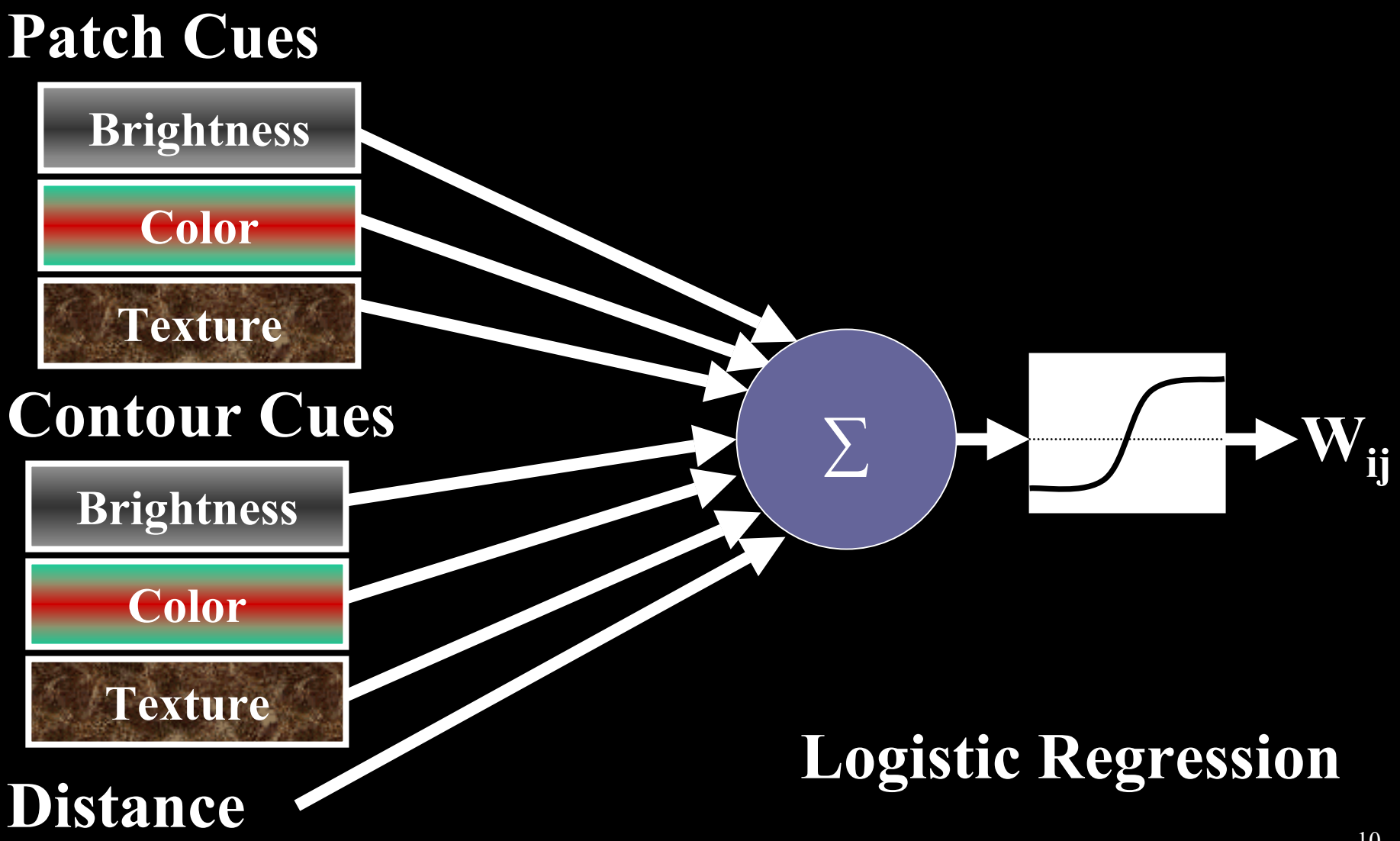
W_{ij} is our estimate of the probability that i and j lie in the same segment

Classifiers for Cue Combination

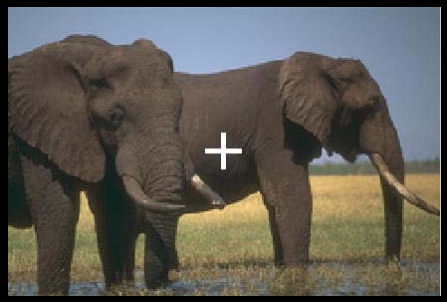
- Classification Trees
  - Top-down splits to maximize entropy, error bounded
- Density Estimation
  - Adaptive bins using k-means
- Logistic Regression, 3 variants
  - Linear and quadratic terms
  - Confidence-rated generalization of AdaBoost (Schapire&Singer)
- Hierarchical Mixtures of Experts (Jordan&Jacobs)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

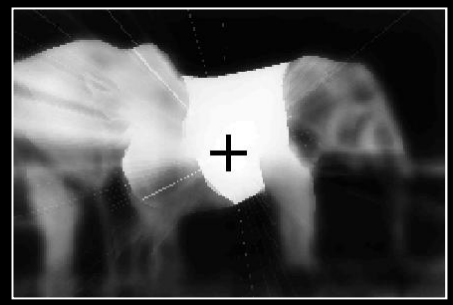
➢ Range over bias, complexity, parametric/non-parametric

Combining Similarity Cues

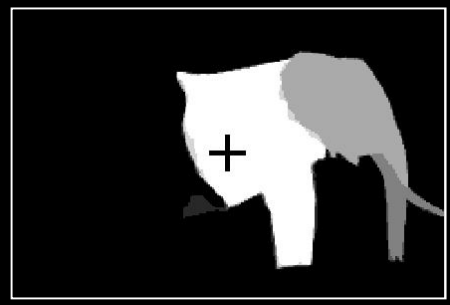


Two Evaluation Methods





Estimate W_{ij}

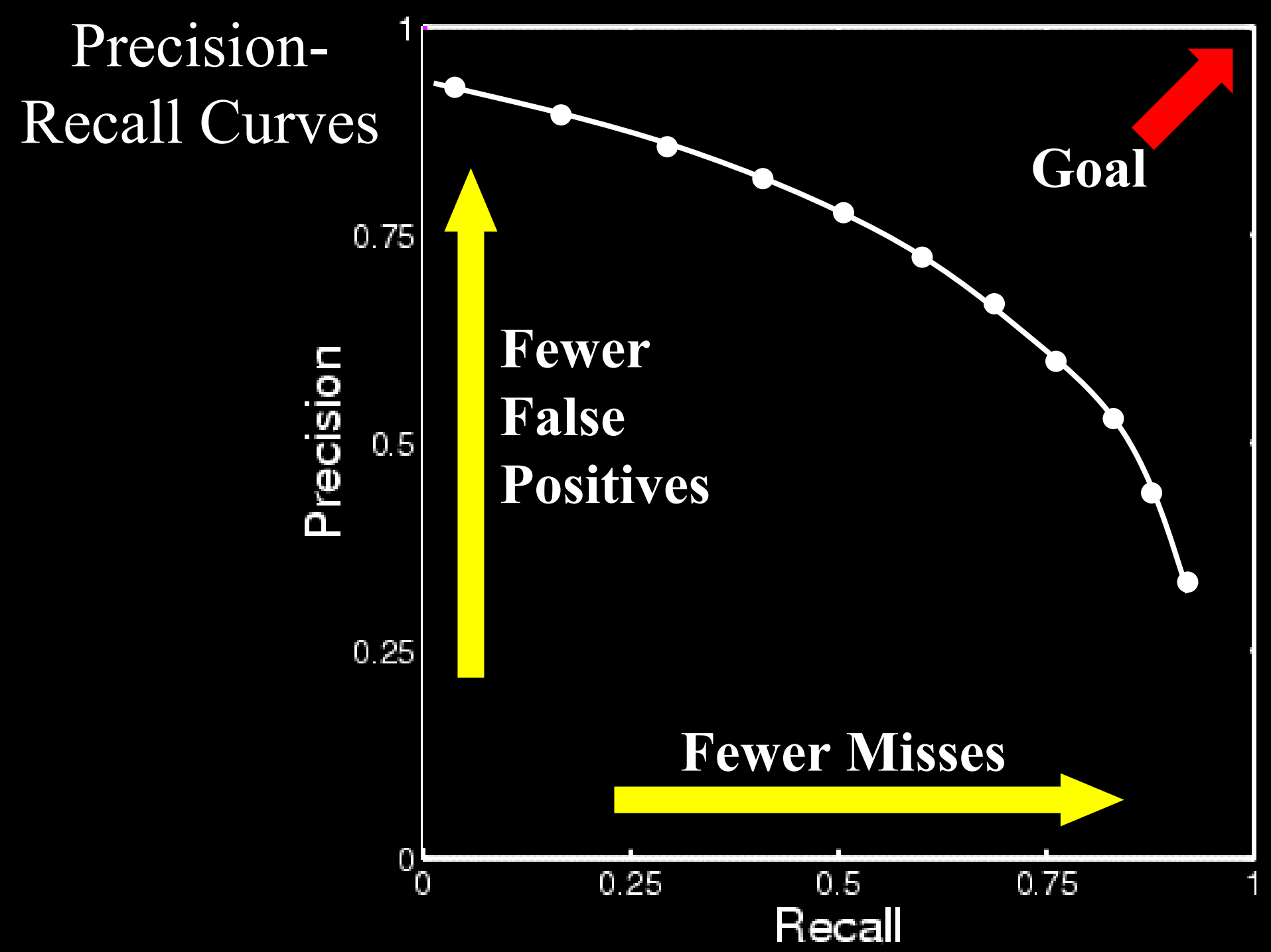


Groundtruth S_{ij}

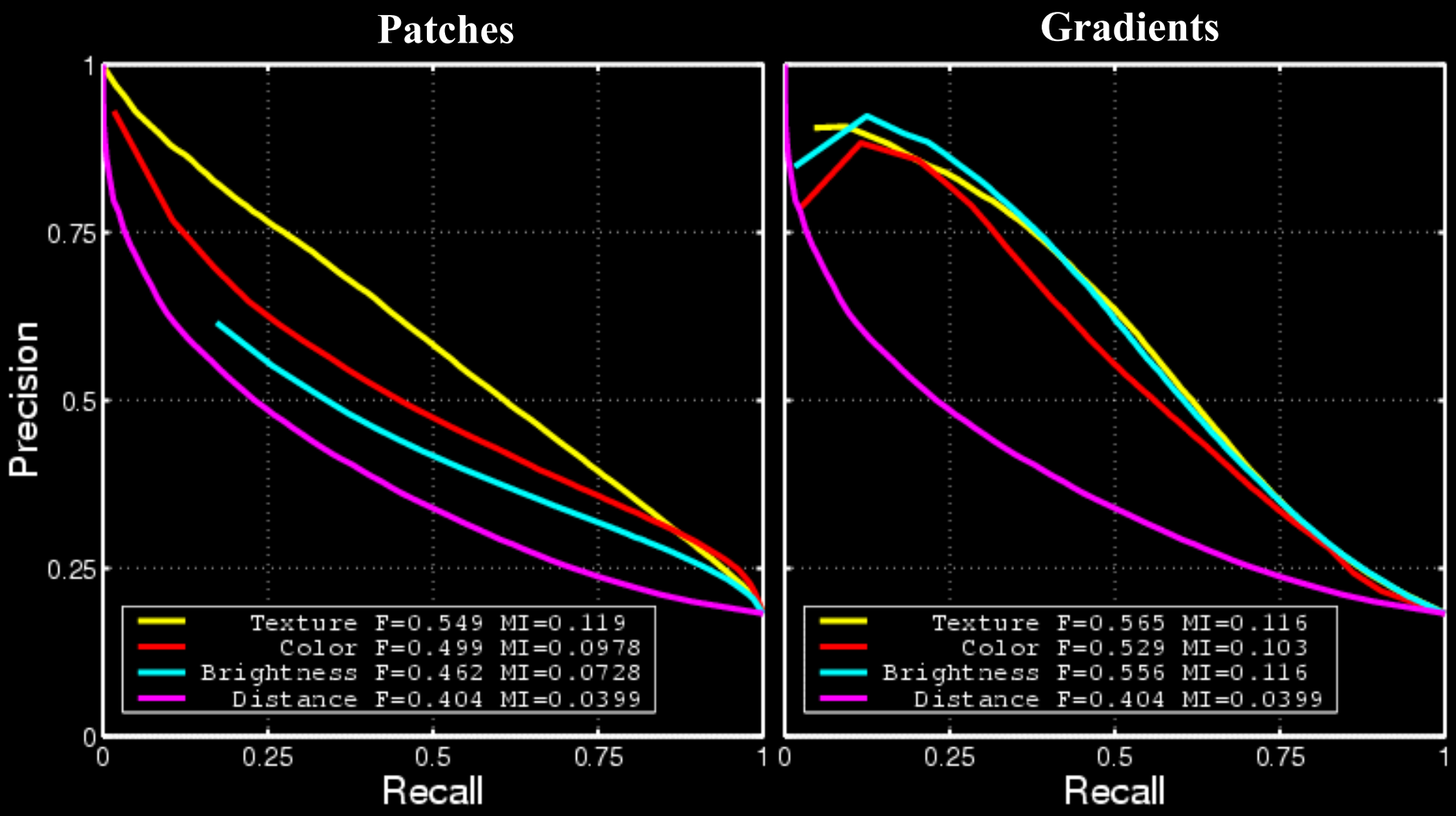
1. Precision-Recall of same-segment pairs
   - Precision is $\mathbf{P}(S_{ij}=1 \mid W_{ij} > t)$
   - Recall is $\mathbf{P}(W_{ij} > t \mid S_{ij} = 1)$
2. Mutual Information between W and S

$$\int p(s,w) \log p(s)p(w) / p(s,w)$$

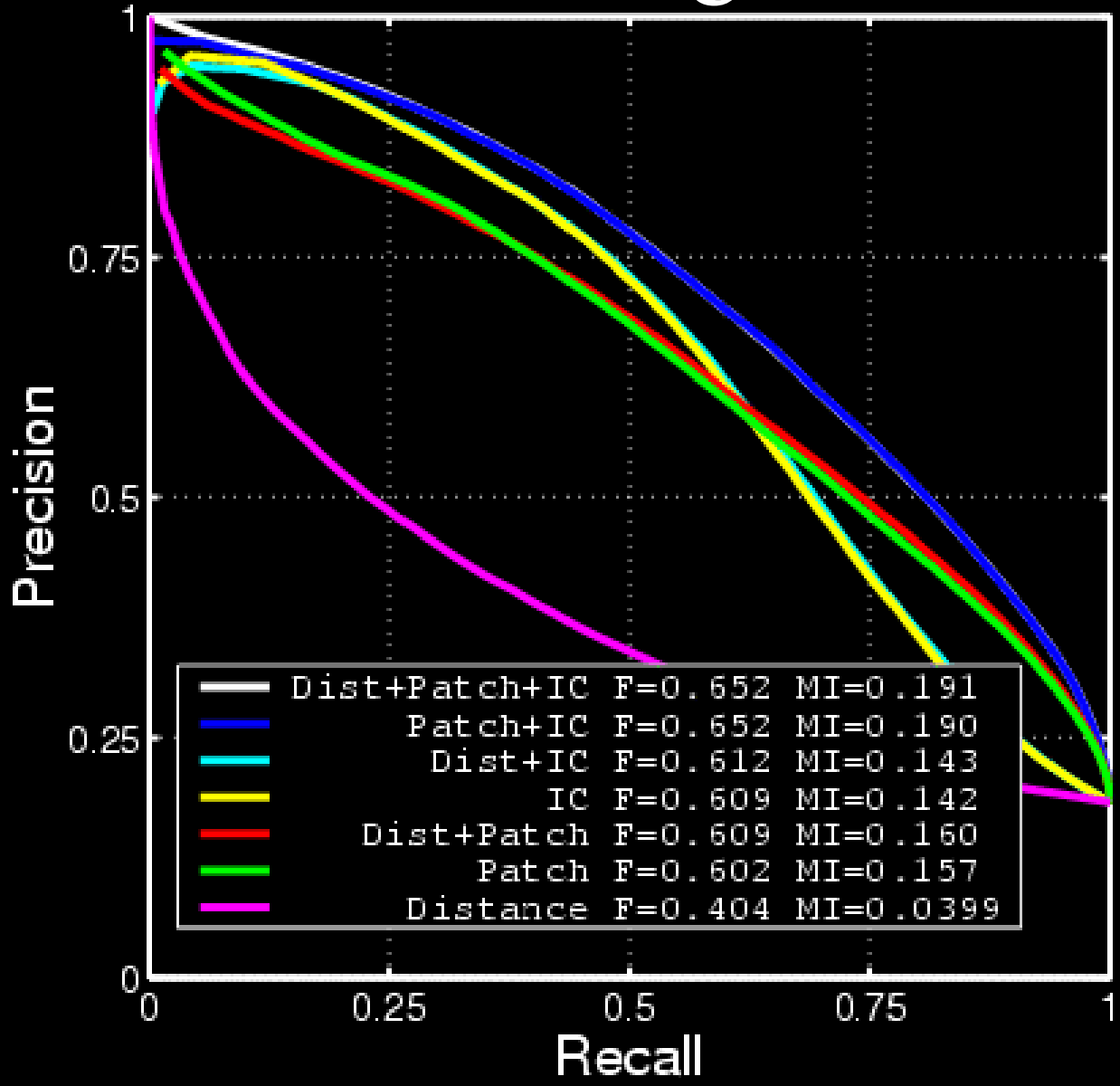
Precision-Recall Curves



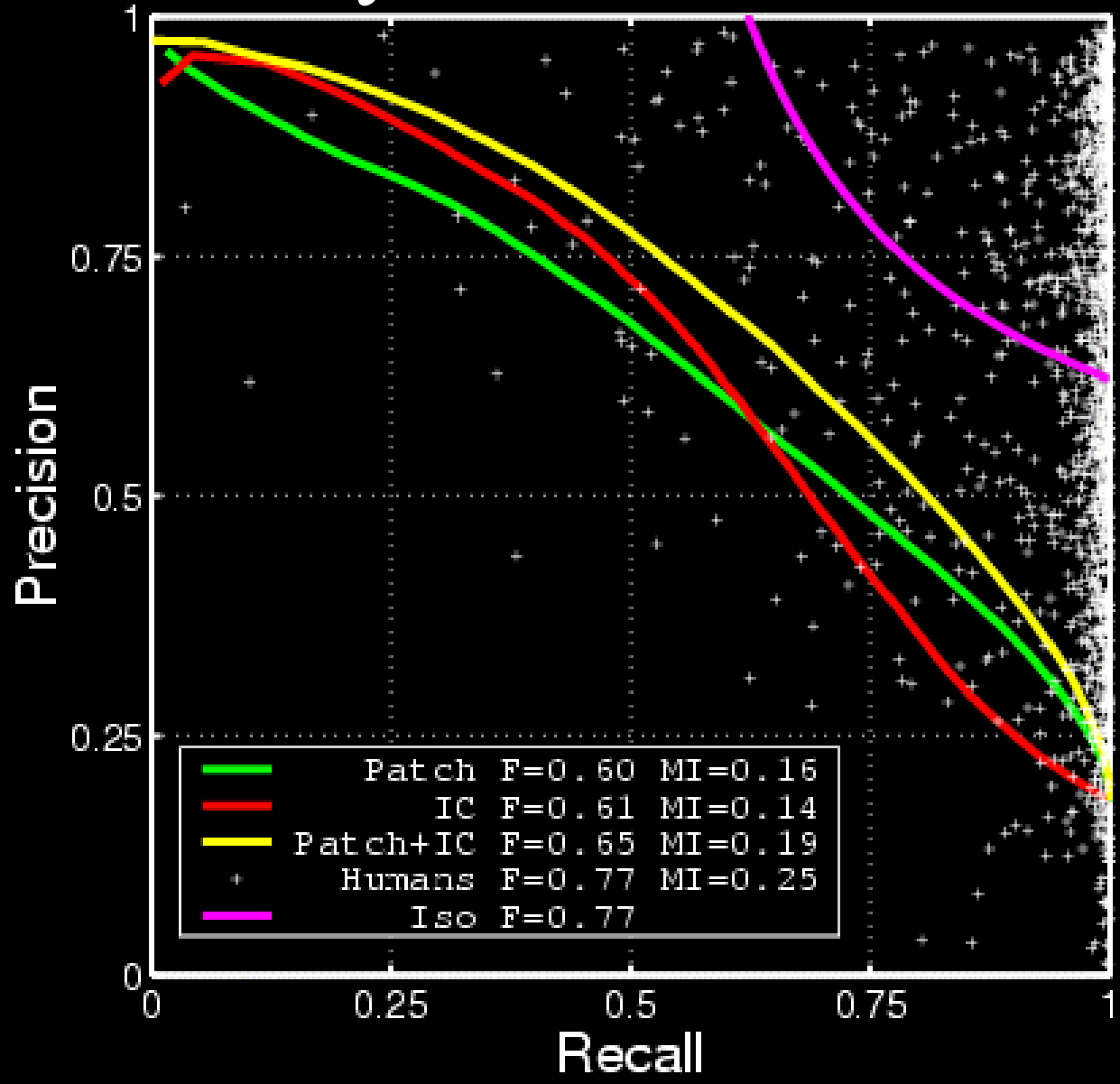
Individual Features



Combining Cues



Affinity Model vs. Humans



Conclusions

- Both Edges and Patches are useful.
- Texture gradients can be quite powerful
- Color patches better than gradients
- Brightness gradients better than patches.
- Proximity is a result, not a cause of grouping

The End

Conclusions

1. Common Wisdom: Use patches only / Use edges only
   Finding : Use both.
2. Common Wisdom : Must use patches for texture
   Finding : Not true, texture gradient is powerful
3. Common Wisdom : Color provides lots of information
   Finding : True, but texture is better
4. Common Wisdom : Brightness patches are a poor cue
   Finding : True, shading and shadows
5. Common Wisdom : Proximity is a grouping cue
   Finding : Proximity is a result, not a cause of grouping

Overview

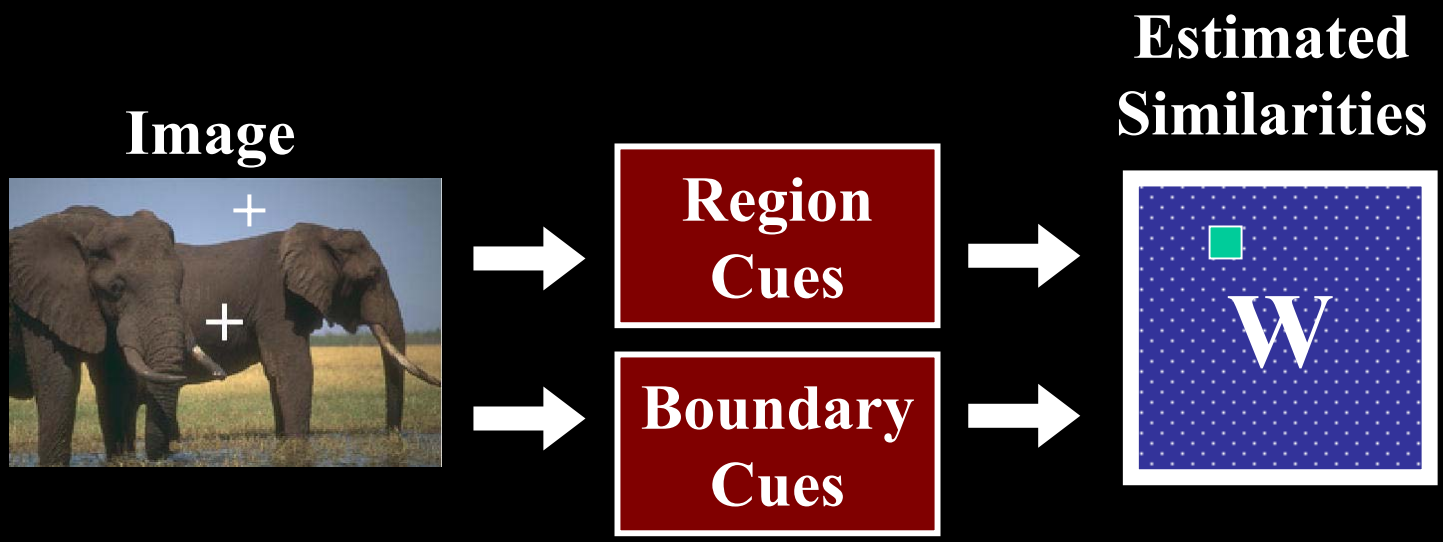
- Gathering the Human Segmentation Dataset
- Computing Similarity Cues
- Modeling Pairwise Similarity
- Evaluating the Relative Power of Similarity Cues

Overview

- Gathering the Human Segmentation Dataset
- Computing Similarity Cues
- Modeling Pairwise Similarity

- Evaluating the Relative Power of Similarity Cues

Pairwise similarity.



- W_{ij} is our estimate of the probability that pixels i and j belong to the same segment given the region and boundary information

Overview

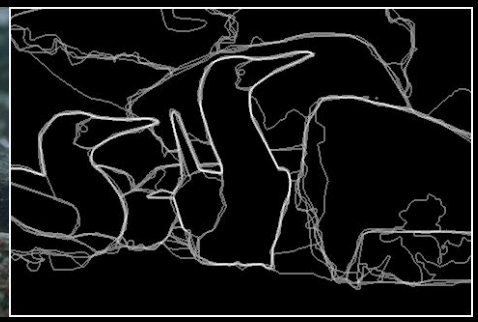
- Gathering the Human Segmentation Dataset
- Computing Similarity Cues
- Modeling Pairwise Similarity

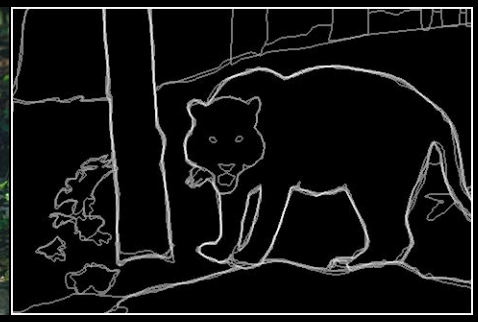
- Evaluating the Relative Power of Similarity Cues

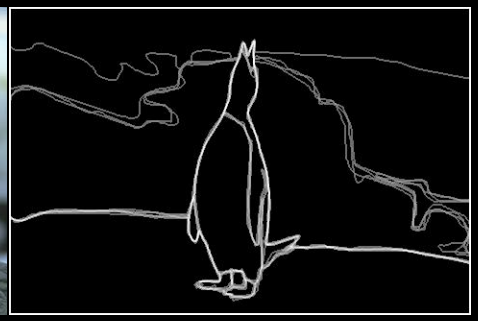
Overview

- Gathering the Human Segmentation Dataset
- Computing Similarity Cues
- Modeling Pairwise Similarity

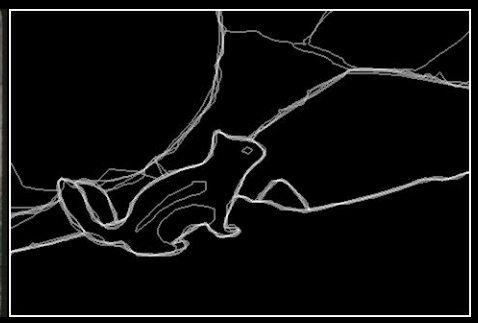
- Evaluating the Relative Power of Similarity Cues

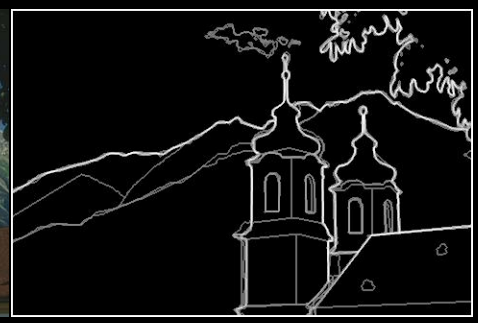





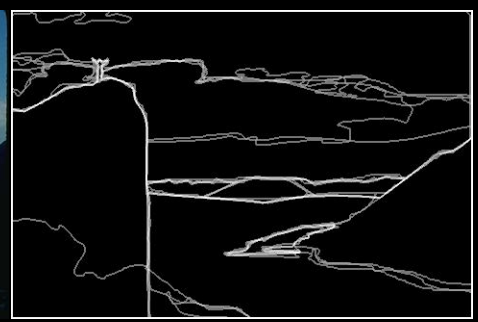












ROC vs. Precision/Recall

Signal \ Truth	P	N
P	TP	FP
N	FN	TN

ROC Curve

Hit Rate = TP / (TP+FN) = 

False Alarm Rate = FP / (FP+TN) = 

PR Curve

Precision = TP / (TP+FP) =  **= Specificity**

Recall = TP / (TP+FN) =  **= Sensitivity**