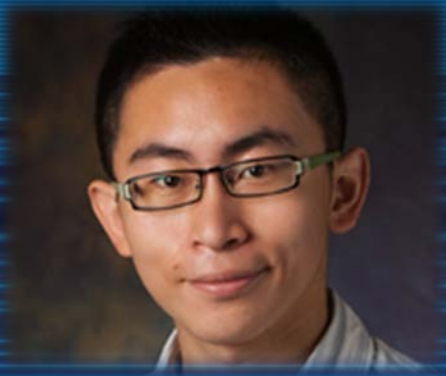
ECE ILLINOIS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Saliency Detection via Divergence Analysis



Jia-Bin Huang

jbhuang1@illinois.edu

Narendra Ahuja

n-ahuja@illinois.edu

Electrical and Computer Engineering
University of Illinois, Urbana-Champaign

What this talk is about?

- The saliency detection problem
- A unifying framework for bottom-up saliency detection algorithms
- Ways to improve the performance

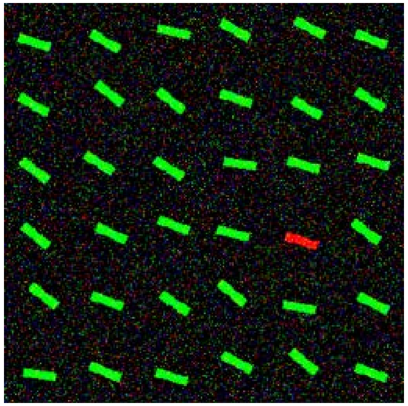

joe-ks

What is Saliency?

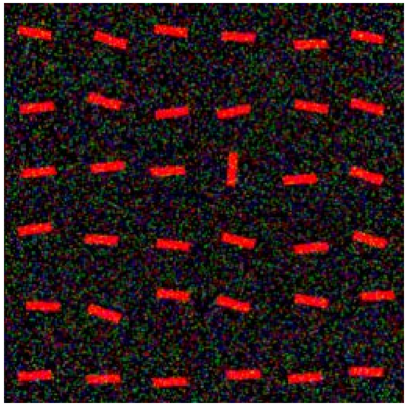
- Visual salience (or visual saliency) is the distinct subjective perceptual quality which makes some items in the world stand out from their neighbors and immediately grab our attention

Where to look?

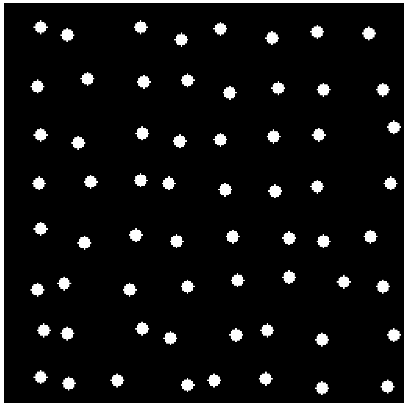
Example of Stimulus



Color



Orientation



Motion

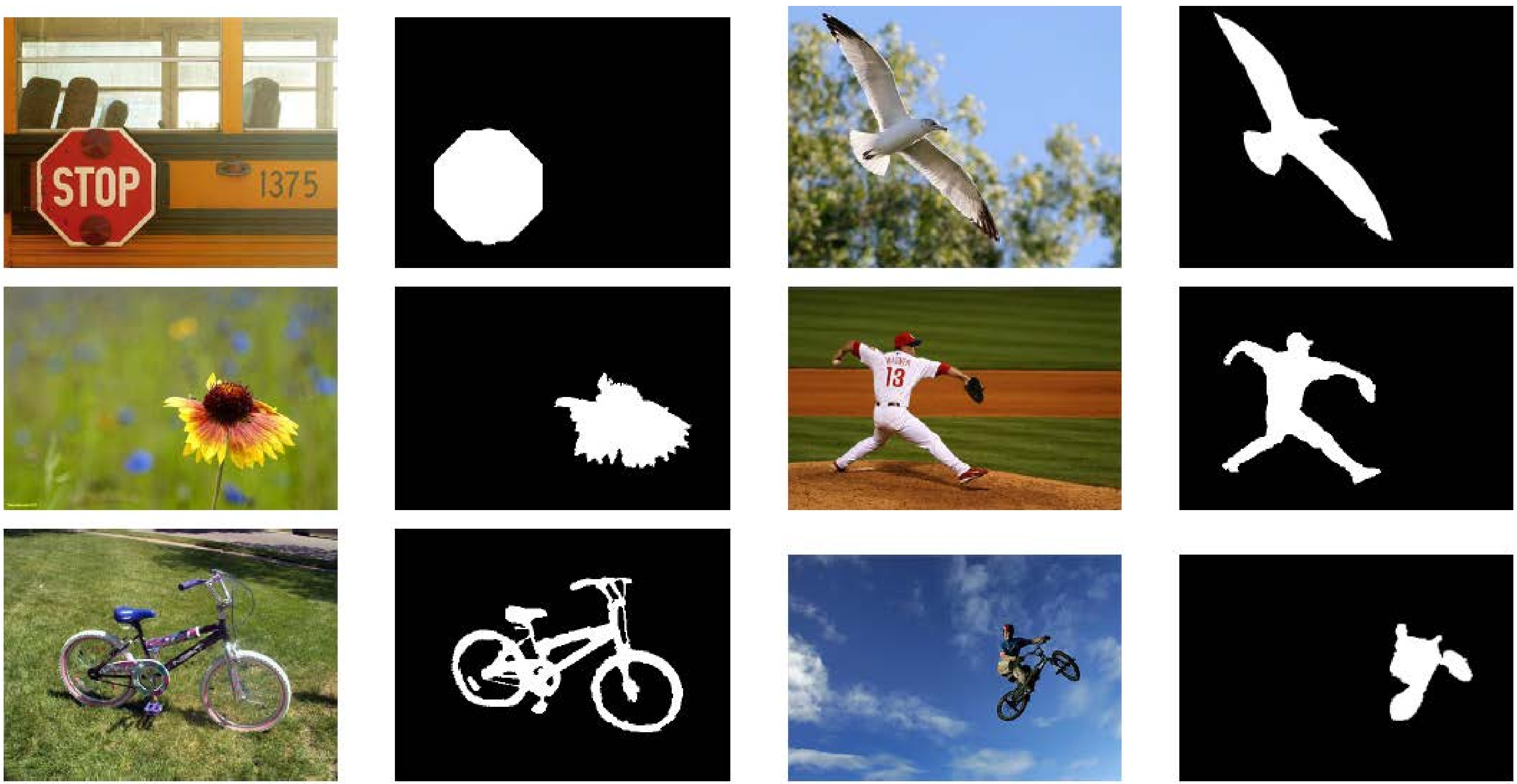
Natural scene

Why Bother?

- Adaptive image compression
- Object-of-interest image segmentation
- Automatic image thumbnail
- (Class-independent) Object detection and recognition
- Visual tracking
- Automatic Image collage
- Content-aware image resizing
- Non-photorealistic rendering
- Understanding mechanism of human visual attention

Problem Setting

- Input: Image -> Output: Saliency map



Design Principles

- **Rarity**
  - [Bruce NIPS 05] [Zhang JOV 08] [Rahtu ECCV 10] [Klein ICCV 11] [Borji CVPR 12]
- **Local complexity**
  - [Kadir IJCV 01]
- **Contrast**
  - [Itti PAMI 98] [Harel NIPS 05] [Ma MM 03] [Achanta CVPR 09] [Achanta ICIP 10] [Cheng CVPR 11] [Goferman CVPR 10] [Perazzi CVPR 12]
- **Spectral**
  - [Hou CVPR 07] [Guo CVPR 08] [Hou PAMI 12] [Li PAMI 12]
- **Learning**
  - [Liu CVPR 07] [Judd ICCV 09] [Oliva ICIP 03] [Torralba Psycho.Rev 06]

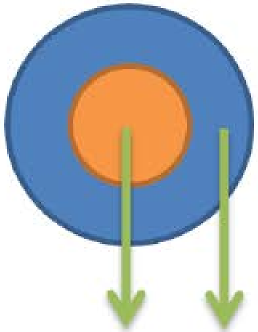

Surprise!
Information!
Complexity!
Rarity!
Spectral!
Contrast!
WTH?
Learning!

Which one is better?

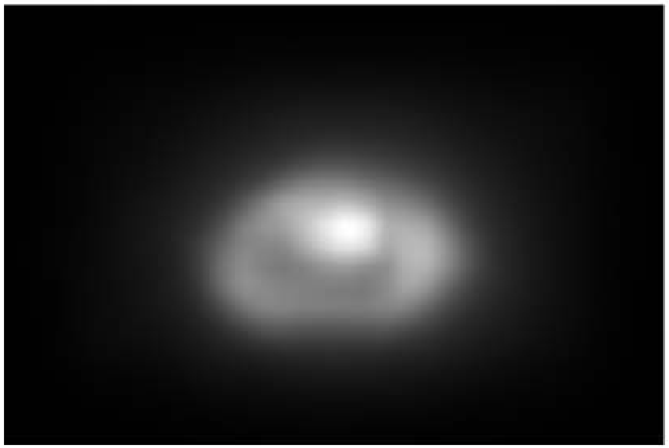
Main Result

- Most of the bottom-up saliency detection algorithms can be rewritten in the form of divergence between probabilistic distributions learned from center and surround

Center-Surround Divergence



$$s_{x_i} = D(P_{C_i} || P_{S_i})$$



- x_i: i_{th} pixel location
- f_{x_i}: feature extracted at x_i (color, texture, motion)
- $\boldsymbol{C_i}$: center support, $\boldsymbol{S_i}$: surround support
- Saliency measure at x_i: $s_{x_i} = D(P_{\boldsymbol{C_i}} || P_{\boldsymbol{S_i}})$

Kullback-Leibler Divergence

- Continuous case
  - $D_{KL}(P||Q) = \int P(x) \log \frac{P(x)}{Q(x)} \mathrm{d}x$
- Discrete case
  - $D_{KL}(P||Q) = \sum_{a \in A} P(a) \log \frac{P(a)}{Q(a)}$
- Has many important operational meanings in detection, estimation and information theory

From Center to Surround

- Assume: $\boldsymbol{C_i} = x_i$
  - $P_{\boldsymbol{C_i}}(f_{x_i}) = 1$
- $D(P_{\boldsymbol{C_i}} || P_{\boldsymbol{S_i}}) = \sum_{f_x} P_{\boldsymbol{C_i}}(f_x) \log \frac{P_{\boldsymbol{C_i}}(f_x)}{P_{\boldsymbol{S_i}}(f_x)}$

  $= -\log P_{\boldsymbol{S_i}}\left(f_{x_i}\right)$ (Shannon's self-information)
- Rarity-based saliency
  - [Bruce NIPS 05] [Zhang JOV 08] [Rahtu ECCV 10] [Klein ICCV 11] [Borji CVPR 12]

From Center to Surround

- Difference of self-information [Rahtu ECCV 10]

$$s_{x_i} = (-\log P_{\boldsymbol{s}_i}(f_{x_i})) - (-\log P_{\boldsymbol{c}_i}(f_{x_i}))$$

$$= \log \frac{P_{\boldsymbol{c}_i}(f_{x_i})}{P_{\boldsymbol{s}_i}(f_{x_i})}$$

- Assume feature channel independence [Klein ICCV 11]
  - $s_{x_i} = \sum_j D_{KL}(P_{\boldsymbol{c}_{i,j}} || P_{\boldsymbol{s}_{i,j}})$
  - $P_{\boldsymbol{c}_{i,j}}$: marginal distribution of j_{th} feature channel.

From Surround to Center

- KL divergence ↔Likelihood theory
  - $D_{KL}(P_{\boldsymbol{S}_i}||P_{\boldsymbol{C}_i}) = \log_{\boldsymbol{n}\to\infty} -\frac{\boldsymbol{1}}{\boldsymbol{n}}\log L\left(f_x|P_{\boldsymbol{C}_i}\right), f_x \sim iid\ \ P_{\boldsymbol{S}_i}$
- Interpretation
  - How well the model learned from C_i can explain samples from S_i
- Contrast-based saliency
  - [Itti PAMI 98] [Harel NIPS 05] [Ma MM 03] [Achanta CVPR 09] [Achanta ICIP 10] [Cheng CVPR 11] [Goferman CVPR 10] [Perazzi CVPR 12]

From Surround to Center

- Assume $P_{\boldsymbol{c_i}}$ follows Laplacian distributions [Zhai ACM MM 06]
  - $s_{x_i} = \sum_{j=1}^{n} |f_{x_i} - f_{x_j}|$

From Surround to Center

- Assume $P_{\boldsymbol{c_i}}$ follows Gaussian distributions
  - $s_{x_i} = \sum_{j=1}^{n} (f_{x_i} - f_{x_j})^2$
- Approximation
  - center surround difference [Itti PAMI 98]
  - mean distance [Achanta CVPR 09]
  - kNN patches [Goferman CVPR 10]
  - high-dimensional Gaussian filters [Perazzi CVPR 12]

Symmetrised Divergence

- Symmetric KL divergence [Borji CVPR 12]
  - $s_{x_i} = D(P_{\boldsymbol{C}_i}||P_{\boldsymbol{S}_i}) + D(P_{\boldsymbol{S}_i}||P_{\boldsymbol{C}_i})$
  - Local and Global Patch Rarities

- λ divergence [Dao NIPS 07]
  - $D_\lambda(P_{\boldsymbol{C}_i}||P_{\boldsymbol{S}_i}) = \lambda\, D_{KL}(P_{\boldsymbol{C}_i}||P_{\boldsymbol{A}_i}) + (1-\lambda)\, D_{KL}(P_{\boldsymbol{S}_i}||P_{\boldsymbol{A}_i})$
  - $P_{\boldsymbol{A}_i} = \lambda\, P_{\boldsymbol{C}_i} + (1-\lambda)\, P_{\boldsymbol{S}_i}$, and $\lambda = |\boldsymbol{C}_i|/|\boldsymbol{A}_i|$
  - Mutual information-based saliency

Symmetrised Divergence

- Cauchy-Schwarz divergence [Cheng CVPR 2011]
  - $D_{CS}(P_{\boldsymbol{c}_i}||P_{\boldsymbol{s}_i}) = -\log \frac{\int P_{\boldsymbol{c}_i}(f_x)P_{\boldsymbol{s}_i}(f_x)\mathrm{d}f_x}{\sqrt{\int P_{\boldsymbol{c}_i}(f_x)^2\mathrm{d}f_x}\sqrt{\int P_{\boldsymbol{s}_i}(f_x)^2\mathrm{d}f_x}}$
  - Estimate $P_{\boldsymbol{c}_i}$ and $P_{\boldsymbol{s}_i}$ with Kernel density estimation
  - Theoretic relations with information theory, graph theory, Mercer kernel and spectral theory. [Jenssen Information Theoretic Learning, 2010]

How to Choose Support?

- Center support $\boldsymbol{C_i}$
  - Single pixel
    - low bias, high variance
  - Patch/window-based
    - Balance bias-variance trade-off. However, hard to determine the optimal size
  - Scale space analysis
    - Scale space extrema or aggregation
  - Region-based
    - Capture potential object boundaries
- Center support $\boldsymbol{S_i}$
  - Notion of local and global saliency

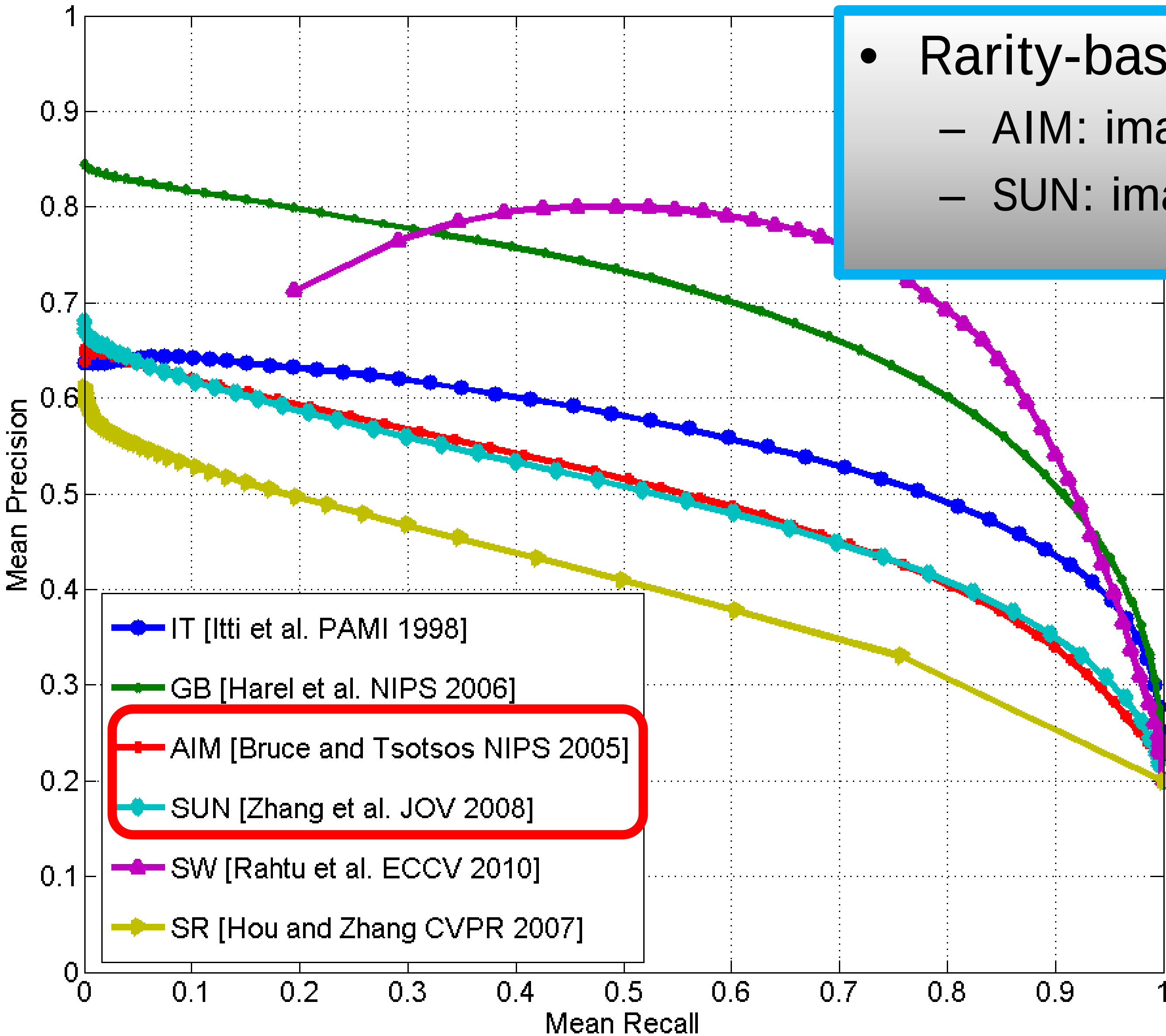
Experimental Results

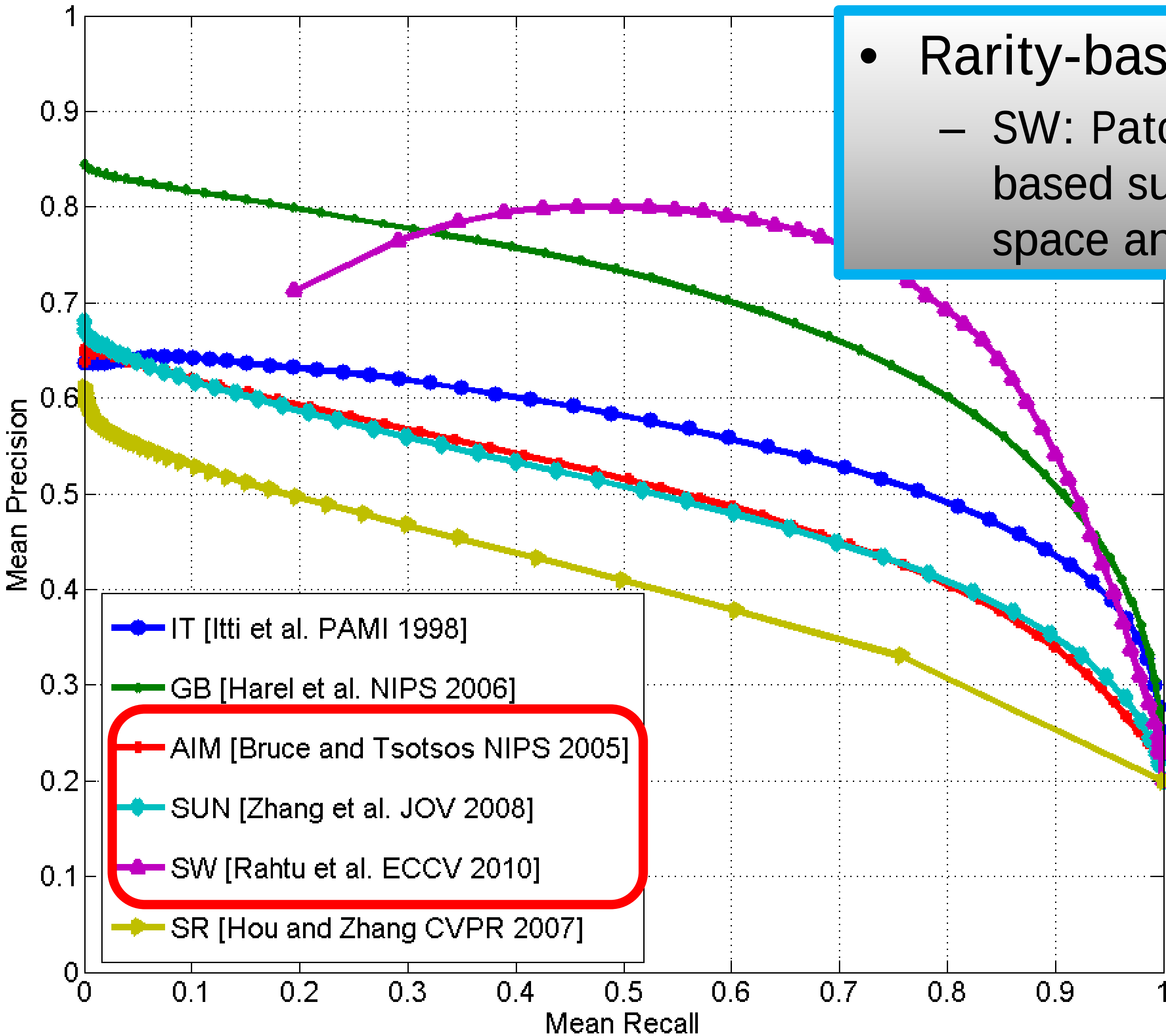
- Datasets
  - MSRA salient object detection dataset
  - 1000 groundtruth binary mask are available from [Achanta CVPR 09]
- Evaluation metric
  - Precision and recall curves

State-of-the-art Saliency Detection Methods

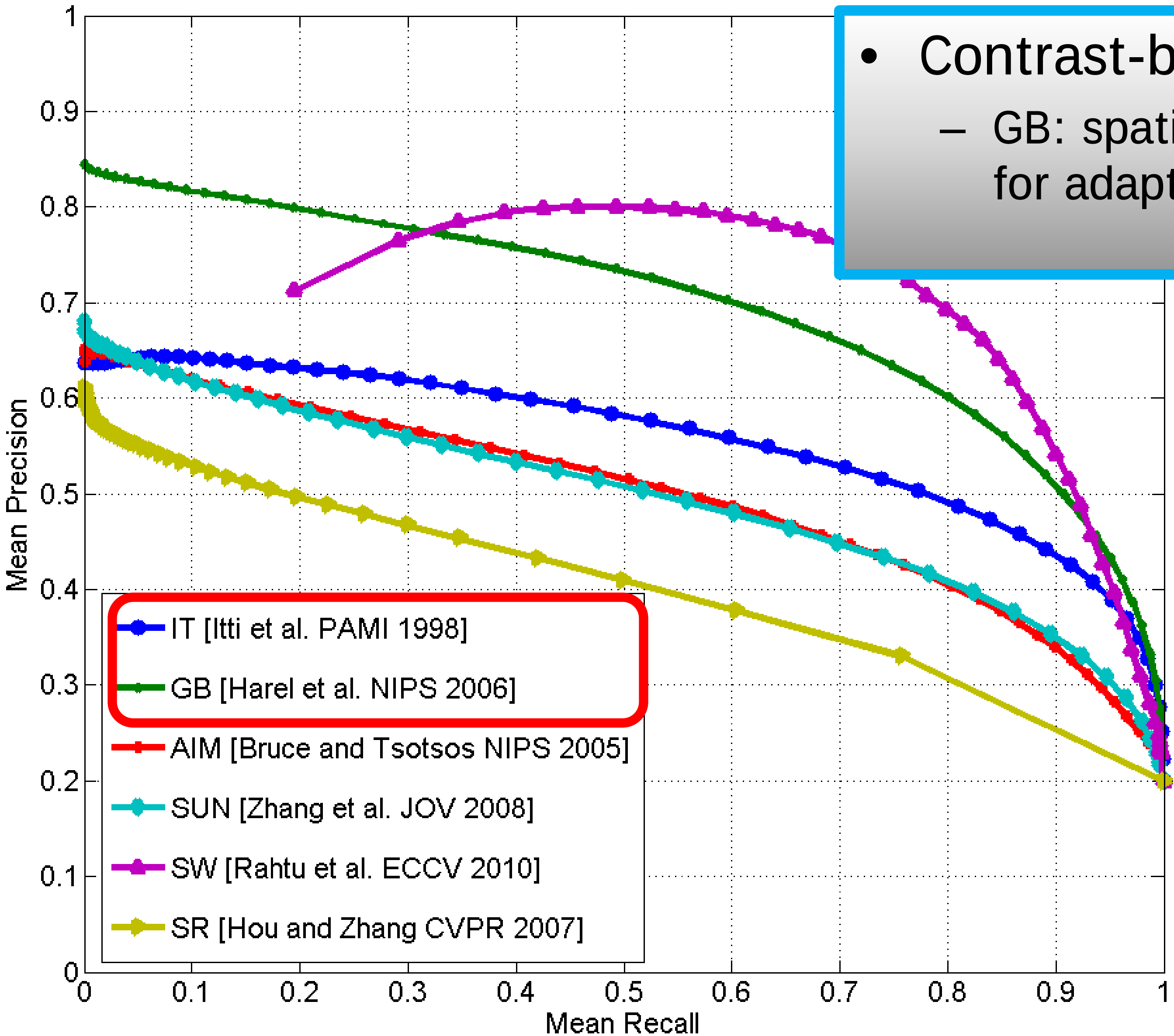
- Center to surround
  - AIM [Bruce NIPS 05] SUN [Zhang JOV 08], SW [Rahtu ECCV 10]
- Surround to center
  - CA [Goferman CVPR 10], AC [Achanta ICIP 10], FT [Achanta CVPR 09], LC [Zhai-ACMMM 06]
- Symmetrised divergence
  - HC, RC, [Cheng CVPR 11], IT [Itti PAMI 98], GB [Harel NIPS 07]
- Spectrum-based
  - SR [Hou CVPR 07]

- Rarity-based
  - AIM: image specific
  - SUN: image indep.

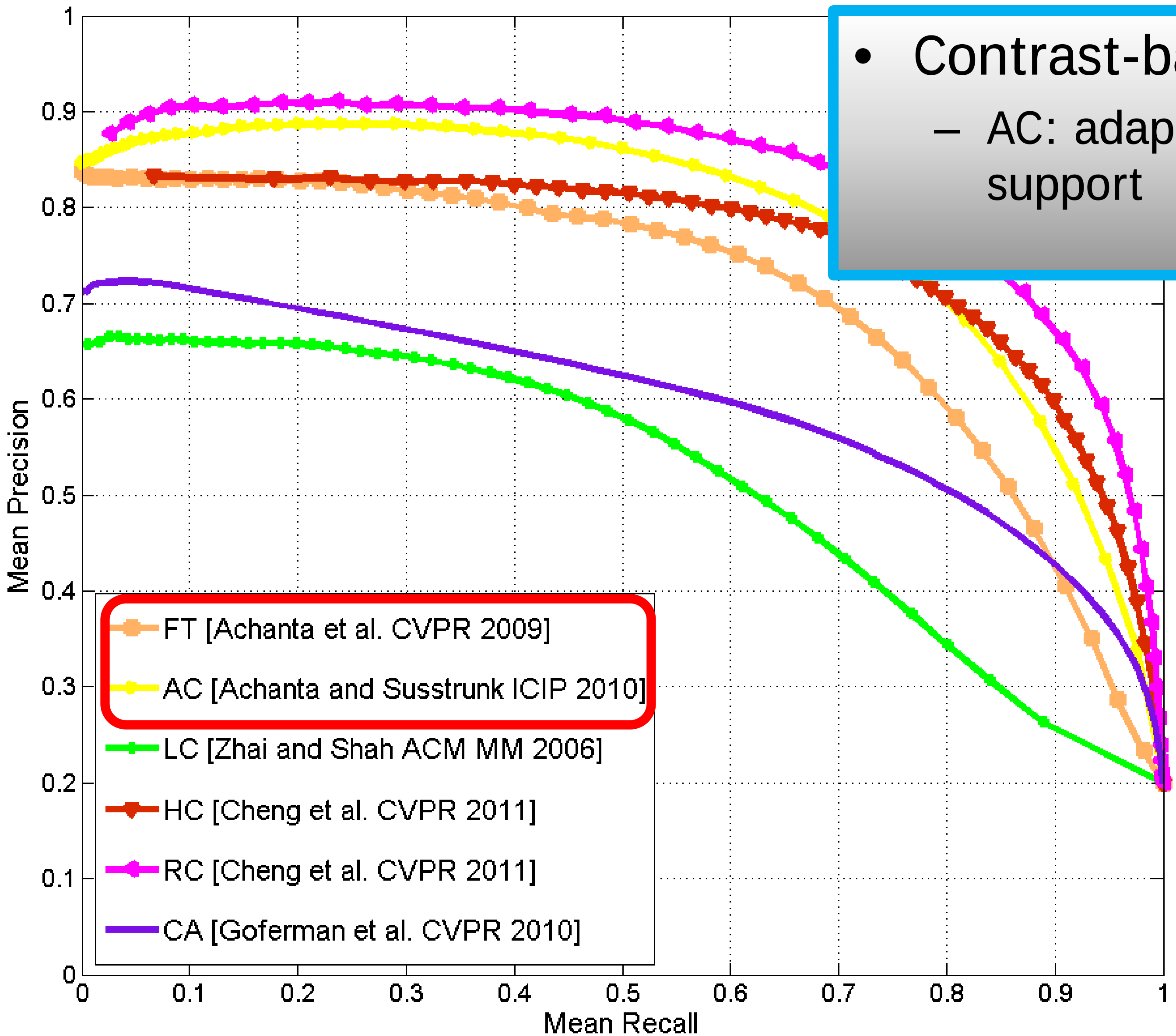




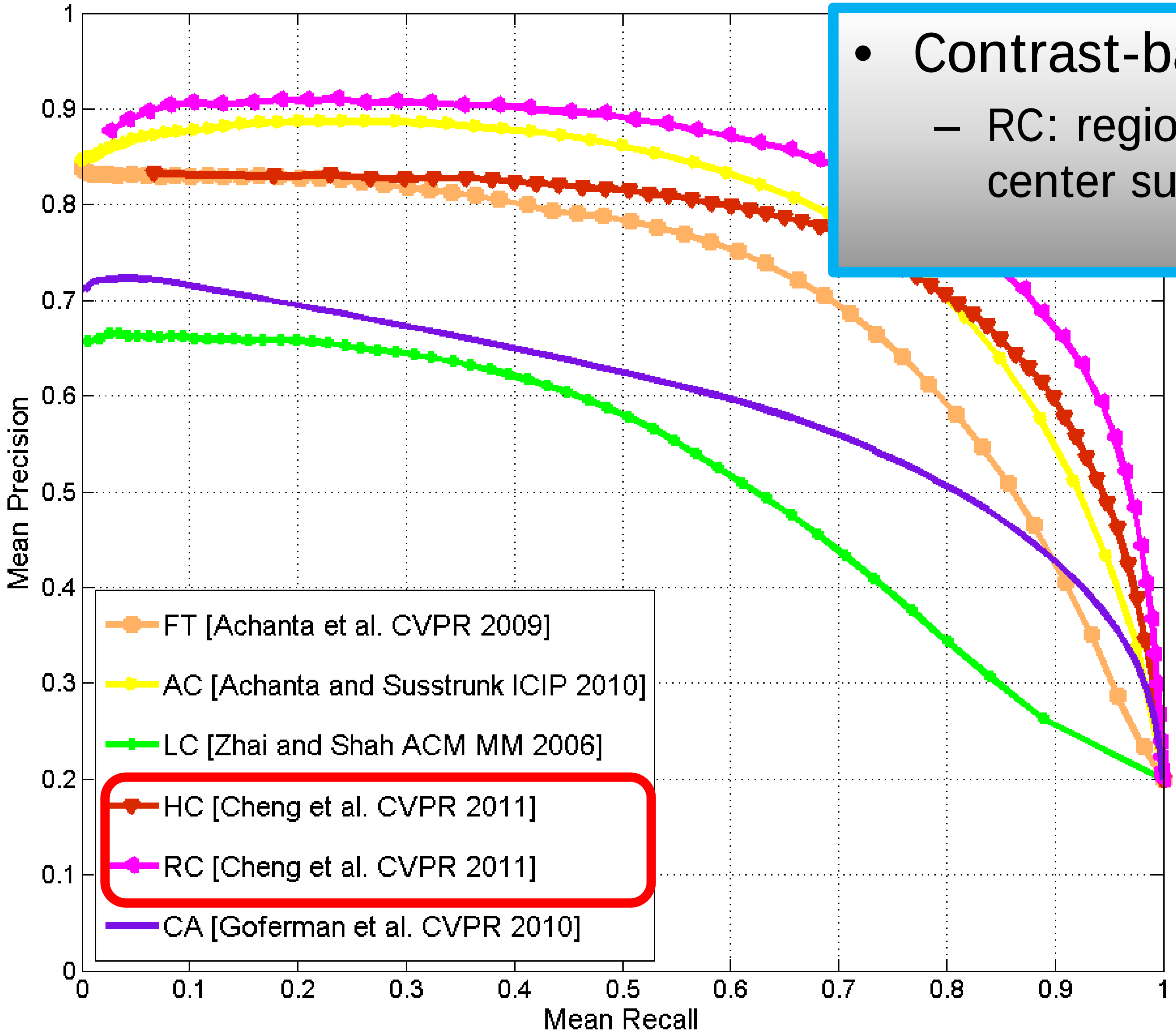
- Rarity-based
  - SW: Patch/window based surround + scale space analysis



- Contrast-based
  - GB: spatial weighting for adaptive surround

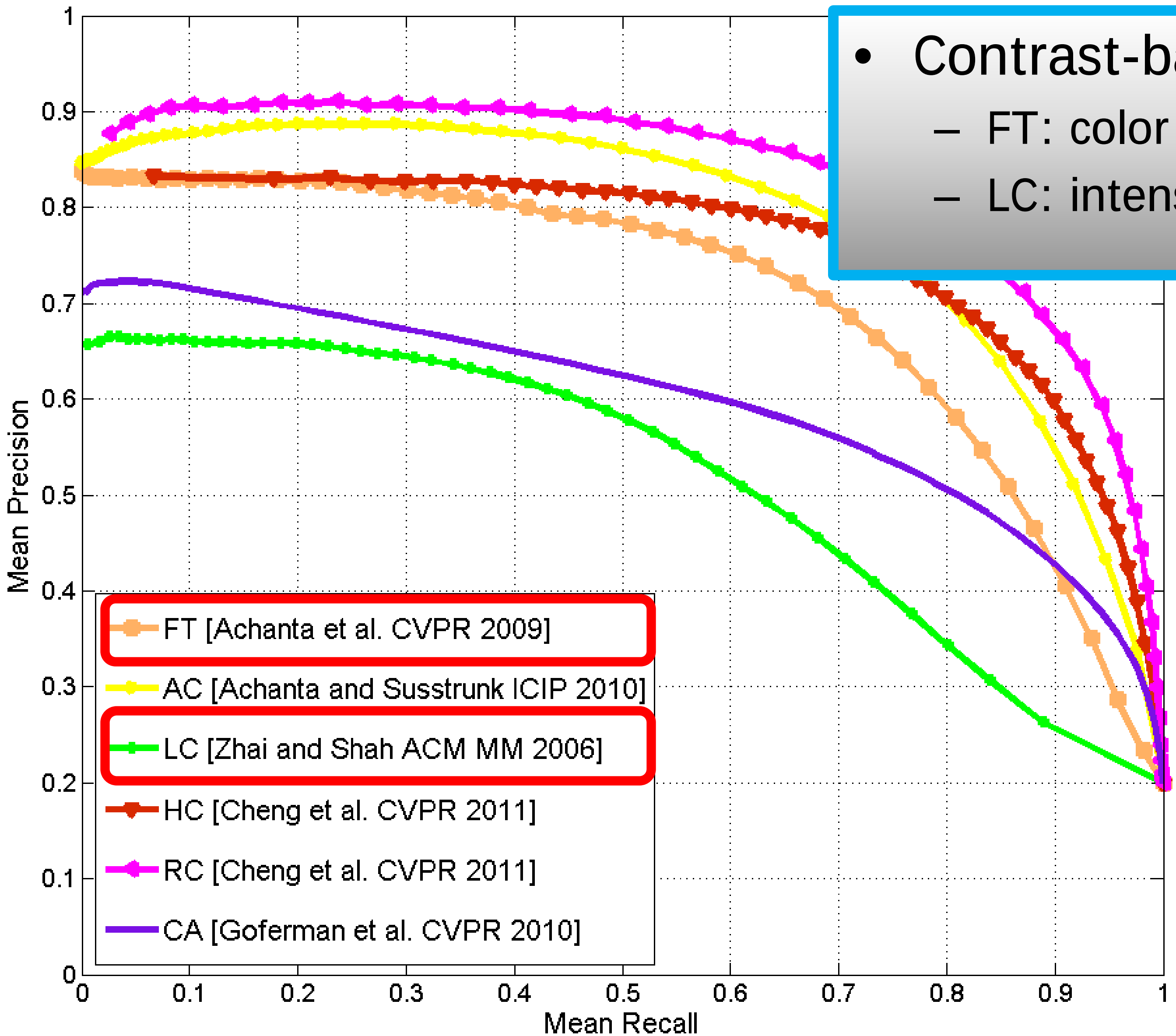


- Contrast-based
  - AC: adaptive surround support



- Contrast-based
  - RC: region-based center support

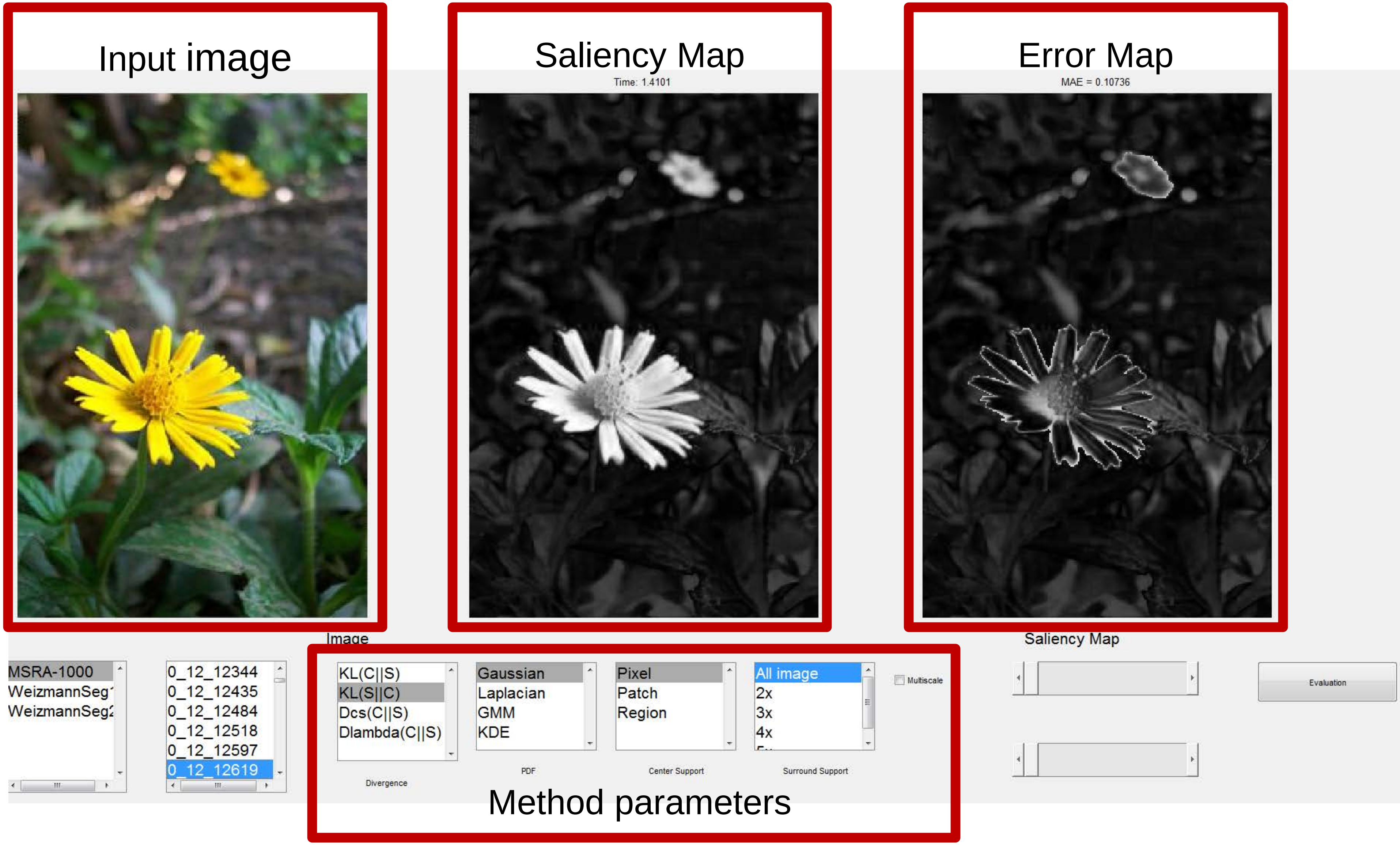
- Contrast-based
  - FT: color feature
  - LC: intensity feature



Lessons Learned

- Most of the bottom-up saliency detection algorithms are in fact close related
  - Not exhaustive, e.g., spectral-based methods
- How to improve the performance?
  - Richer features
  - Less approximation
  - Adaptive center/surround support

Future Work


Input image
Saliency Map
Time: 1.4101
Error Map
MAE = 0.10736
Image
Saliency Map
MSRA-1000
0_12_12344
0_12_12435
0_12_12484
0_12_12518
0_12_12597
0_12_12619
KL(C||S)
KL(S||C)
Dcs(C||S)
Dlambda(C||S)
Gaussian
Laplacian
GMM
KDE
Pixel
Patch
Region
All image
2x
3x
4x
Multiscale
Evaluation
Divergence
PDF
Center Support
Surround Support
Method parameters


Thank You!
:)

References

- Saliency Detection via Divergence Analysis: An Unified Perspective
  J.B. Huang and N. Ahuja. *ICPR, 2012*
- Frequency-tuned salient region detection
  R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk., *CVPR* 2009
- Saliency detection using maximum symmetric surround.
  R. Achanta and S. Susstrunk, ICIP, 2010.
- State-of-the-art in visual attention modeling
  A. Borji and L. Itti. *PAMI* 2012
- Saliency based on information maximization.
  N. Bruce and J. Tsotsos. *NIPS*, 2005
- Global contrast based salient region detection.
  M. M. Cheng, G. X. Zhang, N. J. Mitra, X. Huang, and S. M. Hu. CVPR, 2011
- The discriminant center-surround hypothesis for bottom-up saliency.
- D. Gao, V. Mahadevan, and N. Vasconcelos. *NIPS* 2007

References

- Context-aware saliency detection.

  S. Goferman, L. Zelnik-Manor, and A. Tal. *CVPR*, 2010.
- Graph-based visual saliency.

  J. Harel, C. Koch, and P. Perona. *NIPS*, 2006
- Saliency detection: A spectral residual approach.

  X. Hou and L. Zhang. *CVPR*, 2007
- A model of saliency based visual attention for rapid scene analysis.

  L. Itti, C. Koch, and E. Niebur. *PAMI* 1998
- Learning to predict where humans look.

  T. Judd, K. Ehinger, F. Durand, and A. Torralba. *ICCV*, 2009
- Center-surround divergence of feature statistics for salient object detection.

  D. A. Klein and S. Frintrop. *ICCV*, 2011
- Saliency detection based on frequency and spatial domain analyses

  J. Li, M. Levine, X. An, and H. He. *BMVC* 2011

References

- Learning to detect a salient object
  T. Liu, J. Sun, N. N. Zheng, X. Tang, and H. Y. Shum, *CVPR* 2007
- Contrast-based image attention analysis by using fuzzy growing
  Y. Ma and H. Zhang, *ACM MM* 2003
- Top-down control of visual attention in object detection
  A. Oliva, A. Torralba, M. Castelhano, and J. Henderson. *ICIP* 2003
- Segmenting salient objects from images and videos
  E. Rahtu, J. Kannala, M. Salo, and J. Heikkil., *ECCV* 2010
- Computational versus psychophysical image saliency: A comparative evaluation study, A. Toet. *PAMI* 2011
- Visual attention detection in video sequences using spatiotemporal cues
  Y. Zhai and M. Shah. *ACM MM* 2006
- Sun: A bayesian framework for saliency using natural statistics
  L. Zhang, M. Tong, T. Marks, H. Shan, and G. Cottrell. *JOV* 2008