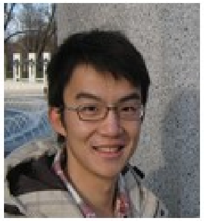
A Crash Course on Visual Saliency Modeling: Behavioral Findings and Computational Models

Location and Dates

Conference on Computer Vision and Pattern Recognition CVPR 2013
The Oregon Convention Center in Portland, Oregon, USA

June 24, 2013, 8:30 - 17:15

Speakers



Ali Borji
University of Southern California (USC)
[primary organizor]

borji@usc.edu

Simone Frintrop
University of Bonn, Germany

frintrop@iai.uni-bonn.de

Laurent Itti
University of Southern California (USC)

itti@usc.edu

Neil D. Bruce
University of Manitoba, Canada

bruce@cs.umanitoba.ca

Xiaodi Hou
California Institute of Technology (Caltech)

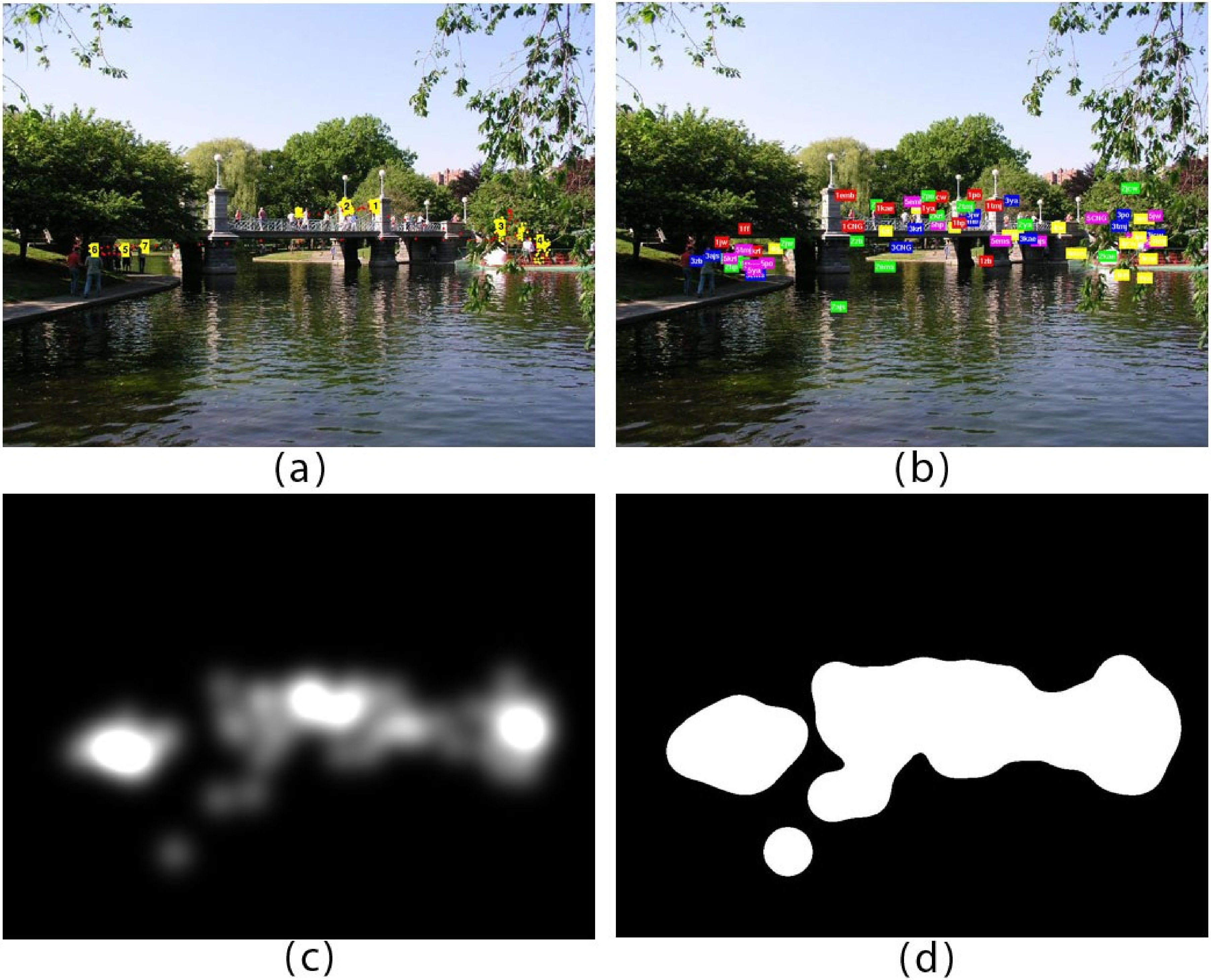
xiaodi.hou@ gmail.com

Schedule

8:30 - 8:45	Introduction to the tutorial	
8:45 - 9:30	Visual attention: Background material	[Ali Borji]
9:30 - 10:15	Attention in daily life	[Ali Borji]
10:15 - 10:30	Break	
10:30 - 11:30	Bayesian and information-theoretic models	[Neil D. Bruce]
11:30 - 12:00	Applications of saliency modeling	[Neil D. Bruce]
12:00 - 13:30	Lunch break	
13:30 - 14:15	Saliency and sparsity	[Xiaodi Hou]
14:15 - 15:00	Towards attentive robots	[Simone Frintrop]
15:00 - 15:30	Attention for 3D object discovery	[Simone Frintrop]
15:30 - 15:45	Break	
15:45 - 16:45	Model comparison and challenges I	[Ali Borji]
16:45 - 17:15	Model comparison and challenges II	[Xiaodi Hou]
17:15 - 18:00	Open forum	

Computational Saliency models, Challenges, Benchmarks, and Future

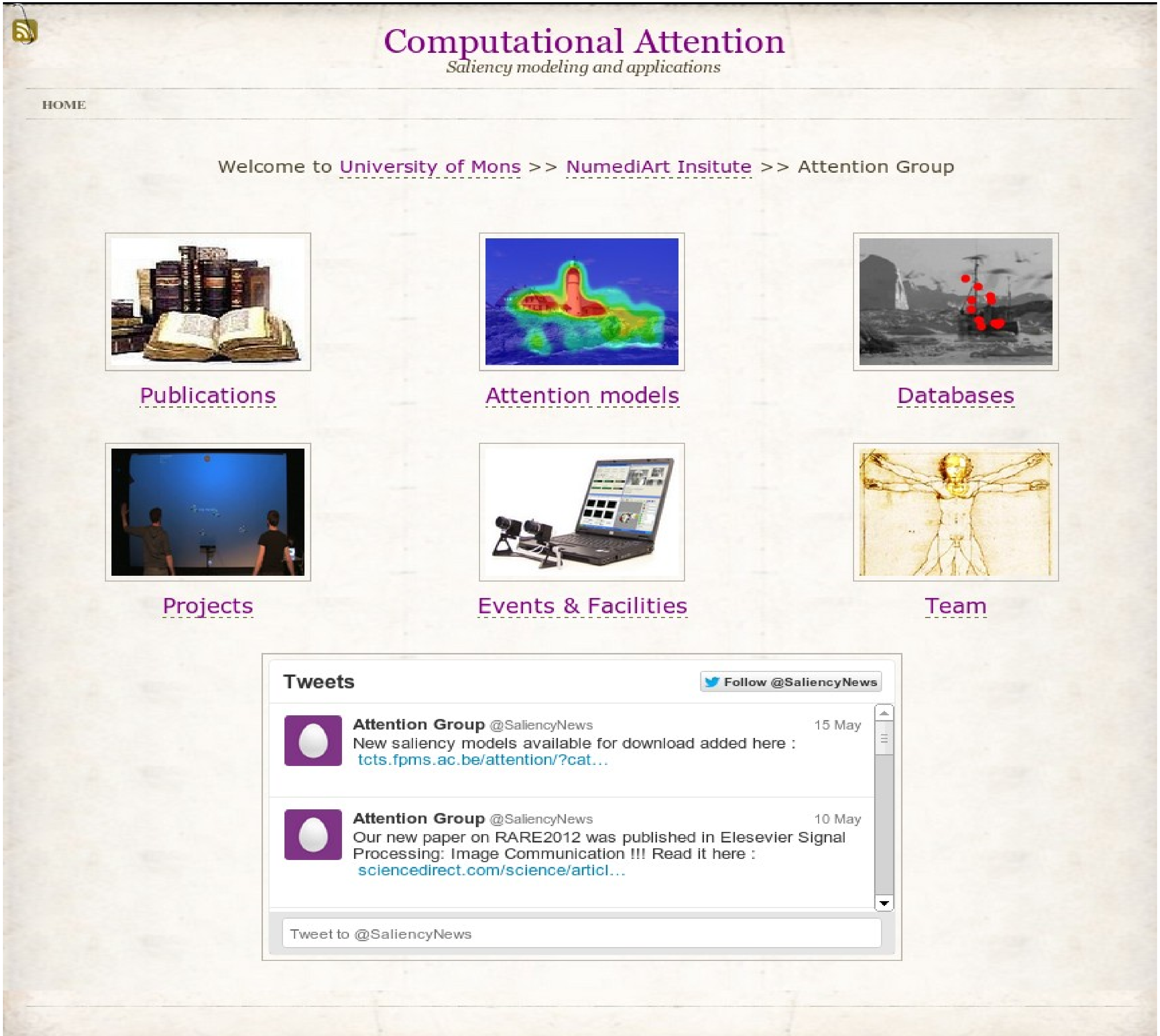
Where do people look?

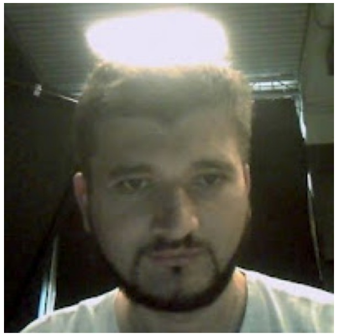


Where do people look? (Judd et al., ICCV 2009)

Plan

- Models
- Challenges
  - Datasets
  - Scores
  - Center-Bias (CB)
- Benchmark
- Summary and Future





Matei Mancas

Roadmap

Matlab Toolbox

Human eye fixations prediction

1. AIM : Neil D. B. Bruce and John K. Tsotsos (2005).
2. STB : Dirk Walther and Christof Koch (2006).
3. GBVS : J. Harel, C. Koch, and P. Perona (2006).
4. SUN : Lingyun Zhang, Matthew H. Tong, Tim K. Marks, Honghao Shan and Garrison W. Cottrell (2008).
5. Dynamic Visual Attention : Xiaodi Hou and Liqing Zhang (2008).
6. WherePeopleLook : Tilke Judd, Krista Ehinger, Frédo Durand and Antonio Torralba (2009).
7. Self-Resemblance Saliency : Hae Jong Seo and Peyman Milanfar (2009).
8. Fast and Efficient Saliency : Rezazadegan Tavakoli H, Rahtu E and Heikkilä J (2011).
9. Saliency Estimation : Naila Murray, Maria Vanrell, Xavier Otazu and C. Alejandro Parraga (2011).
10. Quaternion DCT : B. Schauerte and R. Stiefelhagen (2012) .
11. Top-down Visual Attention : Ali Borji, Dicky N. Sihite, and Laurent Itti (2012) .
12. SignatureSal : Xiaodi Hou, Jonathan Harel and Christof Koch(2012) .

Salient objects detection

1. Saliency Detection: A Spectral Residual Approach : Xiaodi Hou and Liqing Zhang (2007).
2. Detection and Segmentation Saliency : R. Achanta, F. Estrada, P. Wils and S. Süsstrunk (2008).
3. Frequency-tuned Saliency : R. Achanta, S. Hemami, F. Estrada and S. Süsstrunk (2009).
4. Segmenting-based Saliency : Rahtu E, Kannala J, Salo M and Heikkilä J (2010).
5. Saliency-based Image Retargeting : Yuming Fang, Zhenzhong Chen, Weisi Lin, Chia-Wen Lin (2011).
6. Co-saliency Model : Hongliang Li and King Ngi Ngan (2011).
7. Saliency on HVS and Amplitude Spectrum : Yuming Fang, Weisi Lin, Bu-Sung Lee, Chiew Tong Lau, Zhenzhong Chen, Chia-Wen Lin (2012).
8. Wavelet-based Saliency Detection : Nevrez İmamoğlu, Weisi Lin, Yuming Fang (2013).
9. Locally Debiased Region Contrast Saliency : B. Schauerte and R. Stiefelhagen (2013).
10. Saliency Detection Method by Combining Simple Priors : Lin Zhang, Zhongyi Gu and Hongyu Li (2013).

Fixations prediction and objects detection

1. AWS : A. Garcia-Diaz, X. R. Fdez-Vidal, X. M. Pardo, and R. Dosil (2009).
2. Visual Saliency Based on Lossy Coding : Yin Li, Yue Zhou, Junchi Yan, Zhibin Liu, Lei Xu, Xiaochao Yang and Jie Yang (2009).
3. Esaliency : Tamar Avraham and Michael Lindenbaum (2010).
4. Context Aware Saliency : Stas Goferman, Lihi Zelnik-Manor and Ayellet Tal (2010).
5. Frequency and Spatial Saliency : Jian Li, Martin D. Levine, Xiangjing An and Hangen He (2011).
6. Random Center Surround Saliency : T. N. Vikram, M. Tscherepanow and B. Wrede (2012) .
7. Saliency For Image Manipulation : R. Margolin, L. Zelnik-Manor, and A. Tal (2012) .
8. CovSal : Erkut Erdem and Aykut Erdem (2013) .

C/C++ Implementation

Human eye fixations prediction

1. iNVT : Laurent Itti and al. (2001).
2. NMPT : Nicholas Butko and al. (2008).

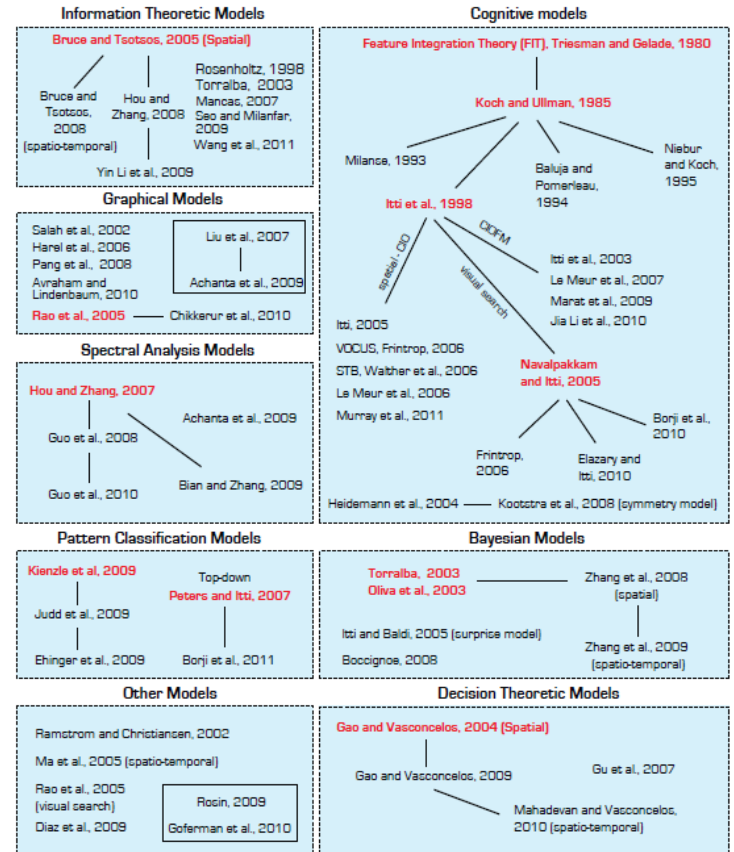
Salient objects detection

1. Salience Distance Transform : P.L. Rosin and G.A.W. West (1995).
2. ADAS : Gert Kootstra and Niklas Bergstrom (2010).
3. Global Contrast Saliency : Ming-Ming Cheng, Guo-Xin Zhang, Niloy J. Mitra, Xiaolei Huang and Shi-Min Hu (2011).
4. Saliency Filters : Federico Perazzi, Philipp Krähenbühl, Yael Pritch and Alexander Hornung (2012).

State-of-the-art in Visual Attention Modeling

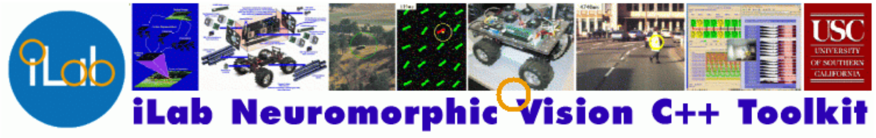
Ali Borji, *Member, IEEE,* and Laurent Itti, *Member, IEEE*

No	Model	Year	f1	f2	f3	f4	f5	f6	f7	f8	f9	f10	f11	f12	f13
Bottom-up (saliency models)															
1	Itti et al. [14]	1998	+	-	-	+	-	-	+	f	+	CIO	C	-	-
2	Privitera & Stark [127]	2000	+	-	-	+	-	-	+	f	+	-	O	-	Stark and Choi
3	Salah et al. [52]	2002	+	+	-	+	-	-	+	-	+	O	G	DR	Digit & Face
4	Itti et al. [119]	2003	+	-	+	+	+	+	+	f	+	CIOFM	C	-	-
5	Torralba [92]	2003	-	+	-	+	-	-	+	s	+	CI	B	DR	Torralba et al.
6	Sun & Fisher [117]	2003	+	-	-	+	-	-	+	-	-	CIO	G	-	-
7	Gao & Vasconcelos [146]	2004	-	+	-	+	-	-	+	s	-	DCT	D	DR	Brodatz, Caltech
8	Ouerhani et al. [210]	2004	+	-	-	+	-	-	+	f	+	CIO+Corner	C	CC	Ouerhani
9	Boccignone & Ferraro [175]	2004	+	-	+	-	+	-	+	f	-	Optical Flow	B	-	BEHAVE
10	Frintrop [50]	2005	+	+	+	+	+	+	+	f/s	+/-	CIOM	C	-	-
11	Itti & Baldi [145]	2005	+	-	+	+	+	+	-	f	+	CIOFM	B	KL, AUC	ORIG-MTV
12	Ma et al. [33]	2005	+	-	+	+	-	-	+	f	+	M*	O	-	-
13	Bruce & Tsotsos [144]	2005	+	-	-	+	-	+	+	f	+	DOG, ICA	I	KL, ROC	Bruce and Tsotsos
14	Navalpakkam & Itti [51]	2006	-	+	-	+	-	+	+	s	+	CIO	C	-	-
15	Zhai & Shah [103]	2006	+	-	+	+	+	-	+	f	+	SIFT	O	-	-
16	Harel et al. [121]	2006	+	-	-	+	-	-	+	f	+	IO	G	AUC	Bruce and Tsotsos
17	Le Meur et al. [41]	2006	+	-	-	+	-	-	+	f	+	LM*	C	CC, KL	Le Meur et al.
18	Walther & Koch [35]	2006	+	-	-	+	-	+	+	f	+/-	CIO	C	-	-
19	Peters & Itti [101]	2007	+	+	+	+	+	-	+	i	+	CIOFM	P	KL, NSS	Peters and Itti
20	Liu et al. [43]	2007	+	-	-	+	-	-	+	f	-	Liu*	G	F-measure	Regional
21	Shic & Scassellati [74]	2007	+	-	+	+	+	-	+	f	+	CIOM	C	ROC	Shic and Scassellati
22	Hou & Zhang [150]	2007	+	-	-	+	-	+	+	f	+	FFT, DCT	S	NSS	DB of Hou and Zhang, 2007
23	Cerf et al. [167]	2007	+	+	-	+	-	+	+	f/s	+	CIO :)	C	AUC	Cerf et al.
24	Le Meur et al. [138]	2007	+	-	+	+	+	-	+	f	+	LM*	C	CC, KL	Le Meur et al.
25	Mancas [152]	2007	+	-	+	+	+	+	+	f	+	CI	I	CC	Le Meur et al.
26	Guo et al. [156]	2008	+	-	-	+	-	-	+	f	+	CIO	D	CC	Self data
27	Zhang et al. [141]	2008	+	-	-	+	-	+	+	f	+	DOG, ICA	B	KL, AUC	Bruce and Tsotsos
28	Hou & Zhang [151]	2008	+	-	+	+	+	-	+	f	+	ICA	I	AUC, KL	Bruce and Tsotsos, ORIG
29	Pang et al. [102]	2008	+	+	+	+	+	-	+	f	+	CIOM	G	NSS	ORIG, Self data
30	Kootstra et al. [136]	2008	+	-	-	+	-	-	+	f	+	Symmetry	C	CC	Kootstra et al.
31	Ban et al. [172]	2008	+	-	+	+	+	-	+	f	+	CIO+SYM	I	-	-
32	Rajashekar et al. [174]	2008	+	-	-	+	-	-	+	f	+	R*	S	CC	Rajashekar et al.
33	Kienzle et al. [165]	2009	+	-	-	+	-	-	+	f	+	I	P	K*	Kienzle et al.
34	Marat et al. [49]	2009	+	-	+	+	+	-	+	f	+	SM*	C	NSS	Marat et al.
35	Judd et al. [166]	2009	+	-	-	+	-	-	+	f	+	J*	P	AUC	Judd et al.
36	Seo & Milanfar [108]	2009	+	-	+	+	+	+	+	f	+	LSK	I	AUC, KL	Bruce and Tsotsos, ORIG
37	Rosin [169]	2009	+	-	-	+	-	-	+	f	+	C+ Edge	O	PR, F-measure	DB of Liu et al, 2007
38	Yin Li et al. [171]	2009	-	+	+	+	+	+	+	s	+	RGB	S	DR	DB of Hou and Zhang, 2007
39	Bian & Zhang [159]	2009	+	-	+	+	+	+	+	f	+	FFT	S	AUC	Bruce and Tsotsos
40	Diaz et al. [160]	2009	+	-	-	+	-	+	+	f	+	CIO	O	AUC	Bruce and Tsotsos
41	Zhang et al. [142]	2009	+	-	+	-	+	-	+	f	+	DOG, ICA	B	KL, AUC	Bruce and Tsotsos
42	Achanta et al. [158]	2009	+	-	-	+	-	-	+	f	+	DOG	S	PR	DB of Liu et al, 2007
43	Gao et al. [147]	2009	+	-	+	+	+	+	+	f	+	CIO	D	AUC	Bruce and Tsotsos
44	Chikkerur et al. [154]	2010	+	+	-	+	-	+	+	f/s	+/-	CIO	B	AUC	Bruce and Tsotsos, Chikkerur
45	Mahadevan & Vasconcelos [106]	2010	+	-	+	-	+	-	+	-	+	I	D	DR, AUC	SVCL background data
46	Avraham & Lindenbaum [153]	2010	+	+	-	+	-	+	+	f/s	+/-	CIO	G	DR, CC	UWGT, Ouerhani et al.
47	Jia Li et al. [133]	2010	-	+	+	+	+	-	+	f	+	CIO	B	AUC	RSD, MTV, ORIG, Peters and Itti
48	Guo et al. [157]	2010	+	-	+	+	+	+	+	f/s	+/-	FFT	S	DR	Self data
49	Borji et al. [89]	2010	-	+	-	+	-	+	+	s	+/-	CIO	O	DR	-
50	Goeferman et al. [46]	2010	+	-	-	+	-	-	+	-	+	C :)	O	AUC	DB of Hou and Zhang, 2007
51	Murray et al. [200]	2011	+	-	-	+	-	-	+	f	+	CIO	C	AUC, KL	Bruce and Tsotsos, Judd et al.
52	Wang et al. [201]	2011	+	-	-	+	-	-	+	f	+	ICA	I	AUC	Self data
Top-down (general attention models)															
53	McCallum [163]	1996	-	+	-	+	-	+	-	i	+	-	R	-	Self data
54	Rao et al. [23]	1996	-	+	-	+	-	-	+	s	+	CIO	O	-	Self data
55	Ramstrom & Christiansen [168]	2002	-	+	-	+	-	-	+	-	+	CI	O	-	-
56	Sprague & Ballard [109]	2003	-	+	+	-	+	+	+	i	-	S*	R	-	-
57	Renninger et al. [94]	2004	-	+	-	+	-	+	-	s	-	Edgelet	I	DR	Self data
58	Navalpakkam & Itti [80]	2005	-	+	-	+	-	+	+	-	+	CIO	C	-	Self data
59	Paletta et al. [164]	2005	-	+	-	+	-	-	+	-	-	SIFT	R	DR	COIL-20, TSG-20
60	Jodogne & Piater [162]	2007	-	+	-	+	-	-	+	i	-	SIFT	R	-	-
61	Butko & Movellan [161]	2009	-	+	+	+	+	+	+	s	-	-	R	-	-
62	Verma & McOwan [214]	2009	+	-	-	+	-	+	-	s	-	CIO	O	-	-
63	Borji et al. [89]	2010	-	+	-	+	-	-	+	i	-	CIO	R	-	-



Borji and Itti, PAMI 2013

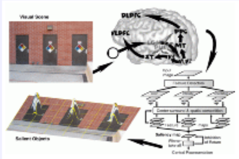
Laurent Itti



- Home
- Overview
- Screenshots
- Documentation
- Downloads
- Publications
- Links

Welcome to the iLab Neuromorphic Vision C++ Toolkit (iNVT)!

The ***iLab Neuromorphic Vision C++ Toolkit*** (iNVT, pronounced ``invent'') is a comprehensive set of C++ classes for the development of neuromorphic models of vision. Neuromorphic models are computational neuroscience algorithr whose architecture and function is closely inspired from biological brains. The iLab Neuromorphic Vision C++ Toolkit comprises not only base classes for images, neurons, and brain areas, but also fully-developed models such as our of bottom-up visual attention and of Bayesian surprise.



Last SVN commit:	Tue Jun 11 14:40:58 2013
Last Forum post:	Wed Jun 12 05:45:22 2013
Last publication:	Tue Feb 12 14:04:58 2013

Features at a glance:

- The source tree is maintained using the Subversion (SVN) revision control system.
- The main development platform is Linux. However, the core programs also compile under Windows (using cygwin) and MacOS X.
- All source code is distributed freely under the GNU General Public License. Registered users get access to our central SVN source code repository and hence receive updates in real-time, not only when we make major releases
- ***Low-level helper classes***, including Point2D, Rectangle, PixRGB<T>, Range, Timer, XWindow, etc.
- ***Template Image<T> and ImageSet<T> classes*** with hundreds of image processing functions and copy-on-write / ref-counting semantics.
- ***Image I/O functions*** for read/write to image files (PNM or PNG) or video streams (various formats).

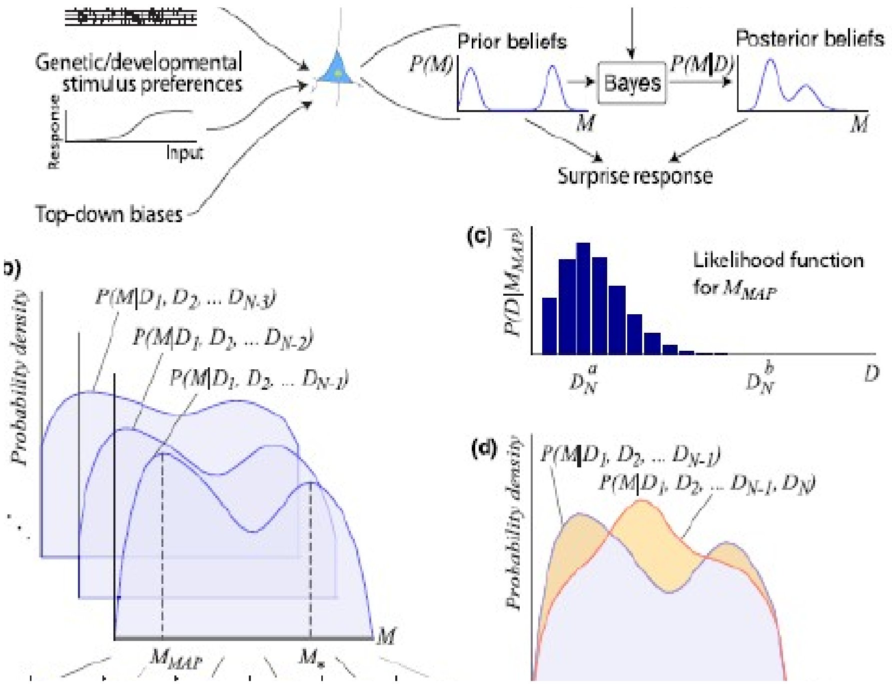
Surprise model

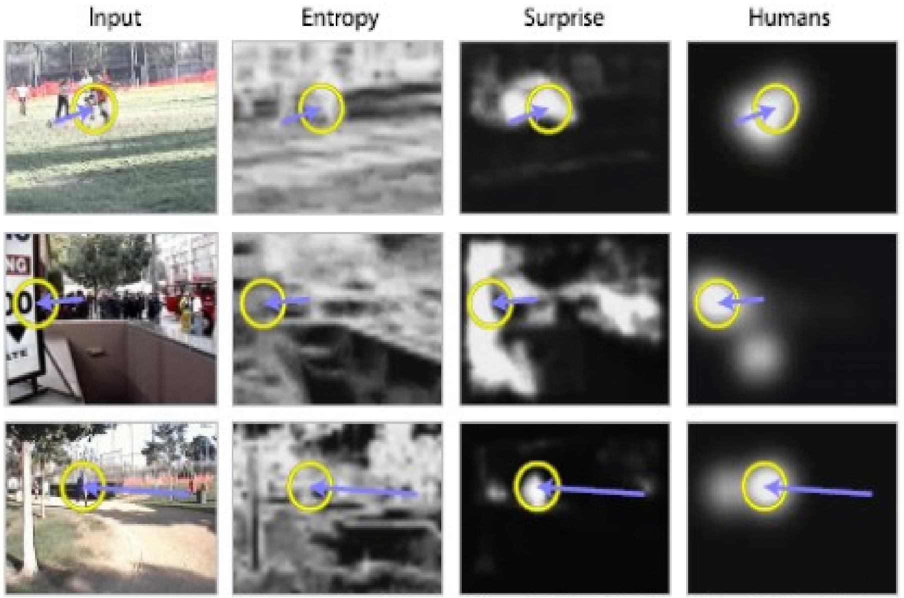
Itti and Baldi, 2006

Laurent Itti

$$\forall M \in \mathcal{M}, \qquad P(M|D) = \frac{P(D|M)}{P(D)} P(M).$$

$$S(D, \mathcal{M}) = KL(P(M|D), P(M)) = \int_{\mathcal{M}} P(M|D) \log \frac{P(M|D)}{P(M)} dM$$

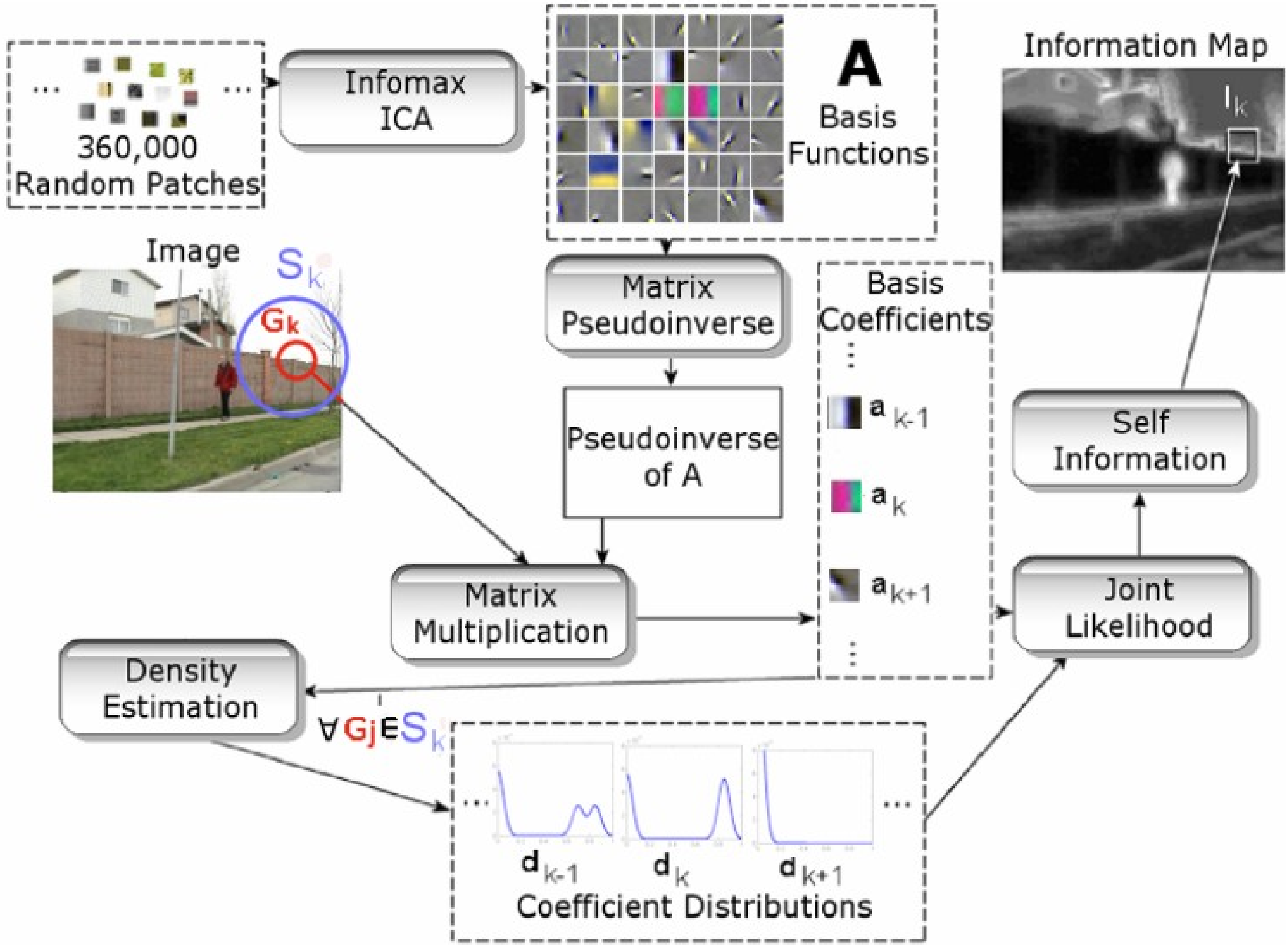




AIM model

Bruce and Tsotsos, 2006

Neil D. Bruce



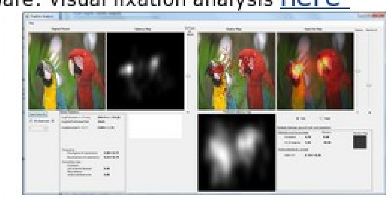
Contact Information

Email: olemeur@irisa.fr
Address: IRISA
Campus Universitaire de Beaulieu
35042 Rennes Cedex - FRANCE
Phone: +33 2.99.84.74.25
Fax: +33 2.99.84.71.71
Project Assistant: +33 2.99.84.72.28
(Huguette Béchu)
Office: F133 BLEU

What's new?

Olivier Le Meur

- Visual dispersion between observers: one paper accepted at ACM Multimedia 2011 (long paper)
- Focal vs ambiant fixations: one paper accepted in iPerception journal 2011
- Open house ESIR demo [here]
- Dynamic saliency map for 3D content: one paper accepted in Cognitive Computation journal 2012
- Video inpainting: one paper accepted at ICIP 2012
- Special session on visual attention SPIE 2012 here NEW !
- One paper accepted at ECCV 2012 (*Super-resolution-based inpainting*, demo Here) NEW !
- One paper accepted in BRM journal (*Similarity metrics for assessing the performance of computational models of visual attention*) NEW ! h
- We release the first version (V1.0) of BRM's software: visual fixation analysis here



NEW !

Recent works on measuring scanpath

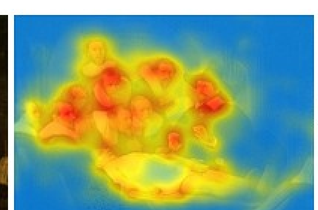
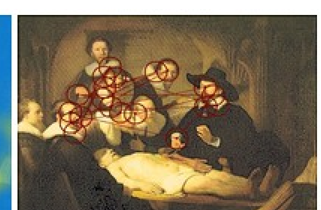
- One paper accepted in Optics Letters (*Saliency detection using regional histograms* [pdf])
- One paper accepted at ICIP'13 (*Memorability of natural scene: the role of attention*) [Here]) NEW !
- One paper accepted in IEEE TIP (*Hierarchical super-resolution-based inpainting*) ([Here]) NEW !
- Supplementary materials SPIE HVEI'13 paper (*How visual attention is modified by disparities and textures changes?*) ([Here])NEW !

Research Areas

- **Visual Attention Understanding And Modelling: 3 Research Axis**
  - Computational modelling of the visual attention:

    From left to right: original picture, heat map and visual scan path (20 first fixation points).

    Leçon d'anatomie, Rembrandt

Graph-based Visual Saliency (GBVS)

Jonathan Harel

[Software] Saliency Map Algorithm : MATLAB Source Code

Below is MATLAB code which computes a salience/saliency map for an image or image sequence/video (either Graph-Based Visual Saliency (GBVS) *or the standard Itti, Koch, Niebur PAMI 1998 saliency map*). See the included readme file for details. I also have a newer, simpler version implementing only the Itti algorithm (see simpsal/readme.txt). Additionally, there is also code to compute a saliency map based on the "Image Signature" as described in this PAMI paper by Xiaodi Hou, Jonathan Harel, and Christof Koch.

Please email me if you have any questions!

Last updated July 24, 2012.

Download:

Please select one of the following to download:

1. [gbvs.zip]
   This package includes an implementation of the full **GBVS algorithm**. It also lets you compu **saliency map**. It includes a function for computing the **ROC score** between eye-movements (points) and a saliency map, and a function for displaying a saliency map overlayed on top of a screenshot below).

   Right after you download the zip file, you must change into the gbvs/ directory and run:

   ```
   >> gbvs_install
   ```

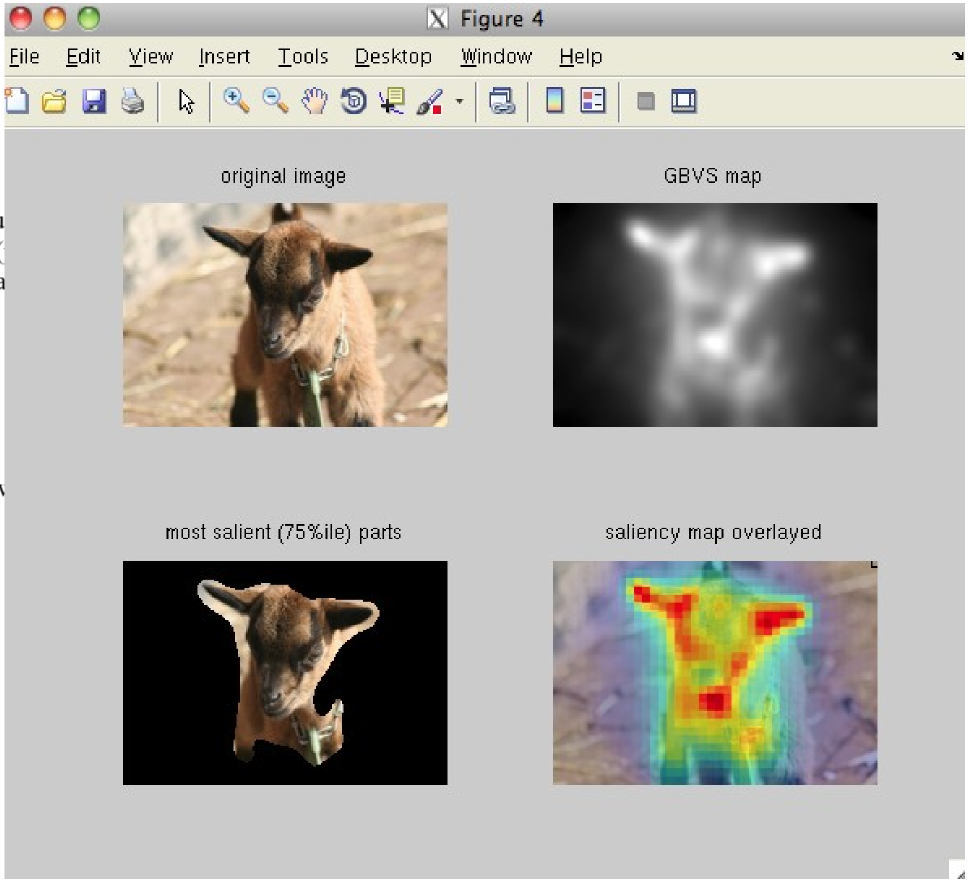
   You only need to run that the first time. Afterwards, you can generate a saliency map as follov

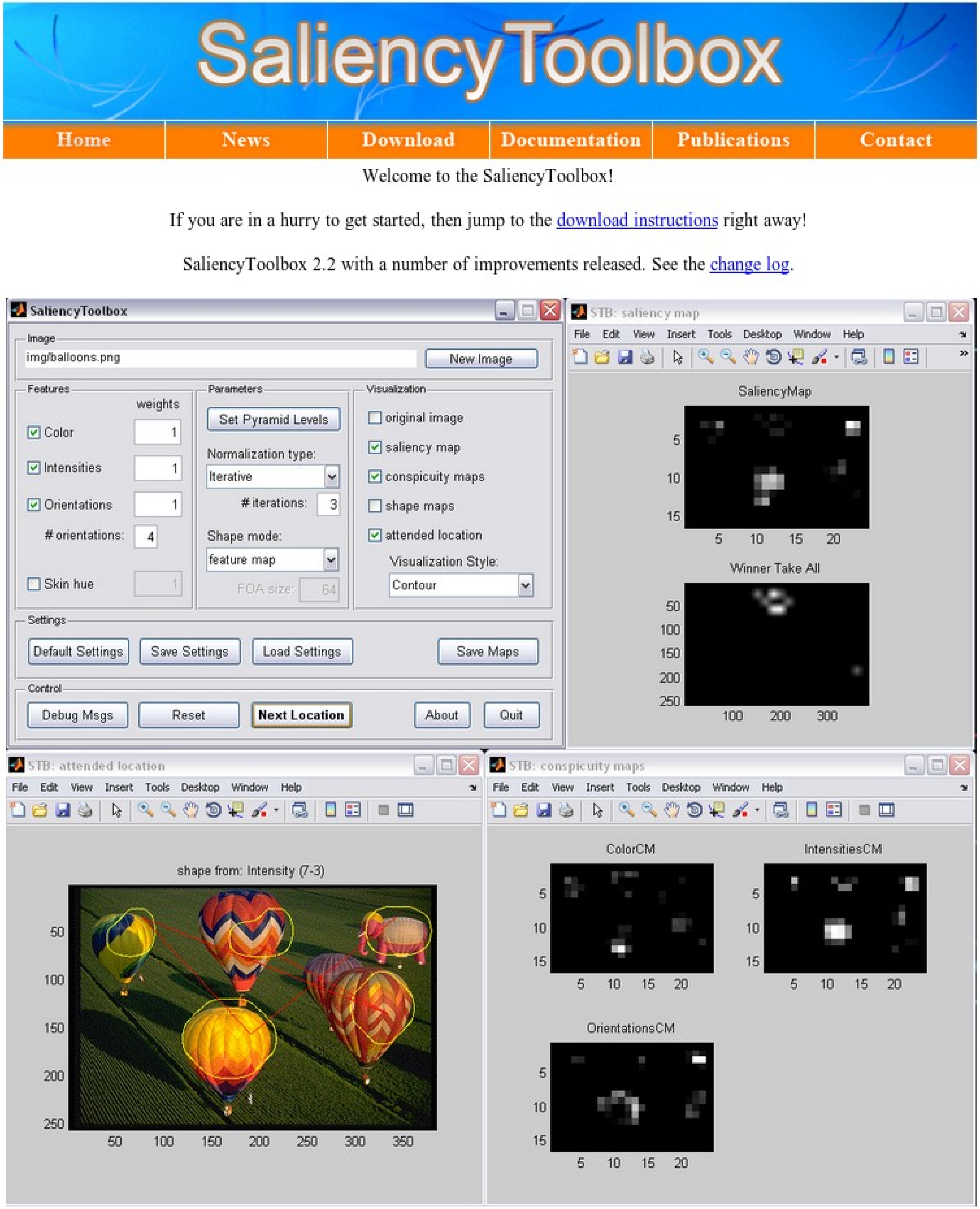
   To load an image:

   ```
   >> img = imread('samplepics/1.jpg');
   ```

   To compute a GBVS map:

   ```
   >> map = gbvs(img); % map.master_map contains the actual saliency map
   ```





Dirk B. Walther

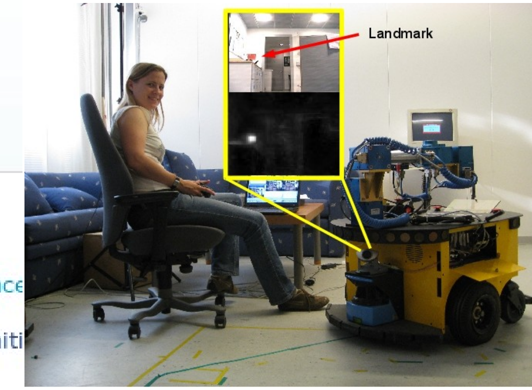
IVS intelligent vision systems

VOCUS

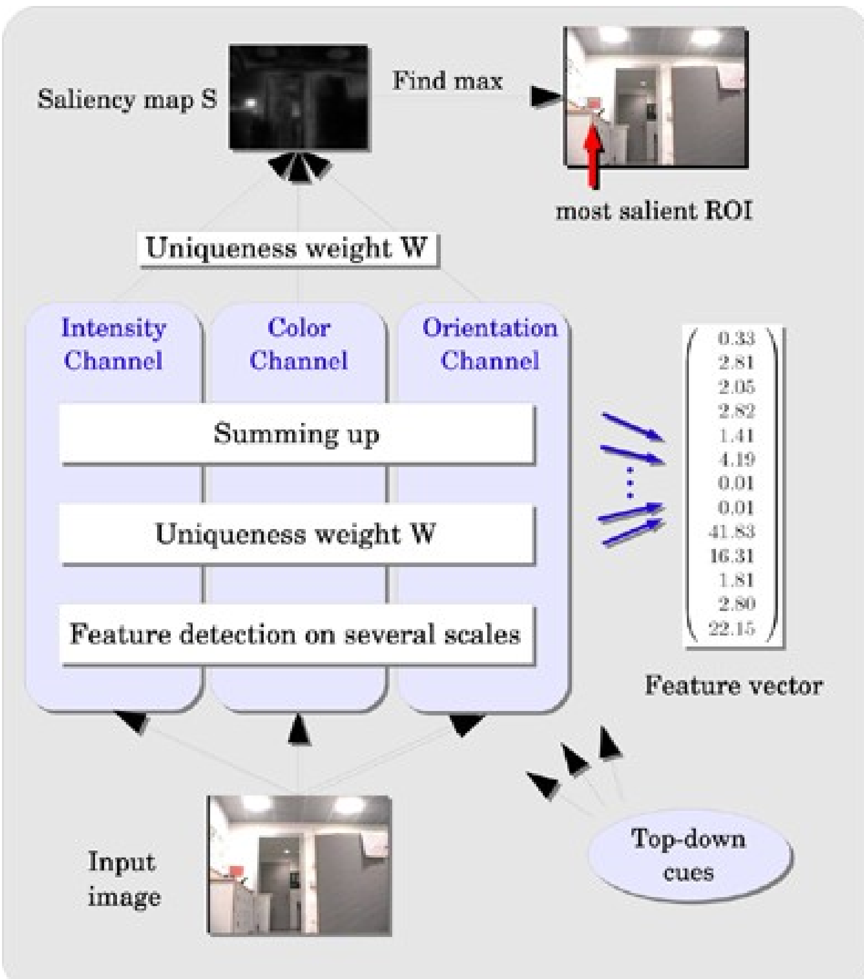
Welcome,

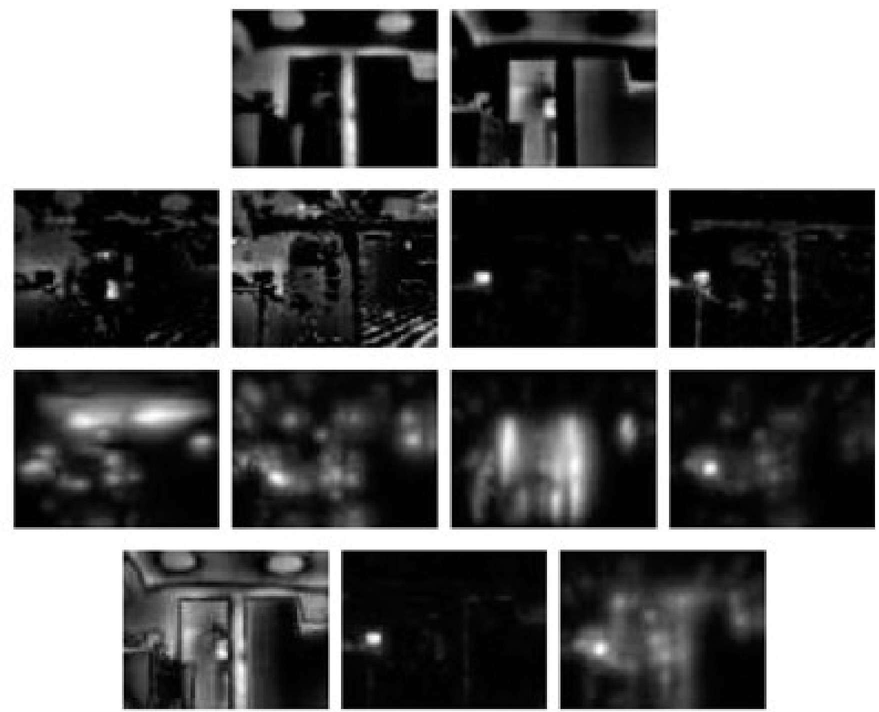
to the Intelligent Vision Systems group directed by Prof. Dr. Armin B. Cremers. The IVS group is part of the Computer Science

Our group consists of two subgroups, the Applied Computer Vision group, headed by PD. Dr. Volker Steinhage, and the Cogniti



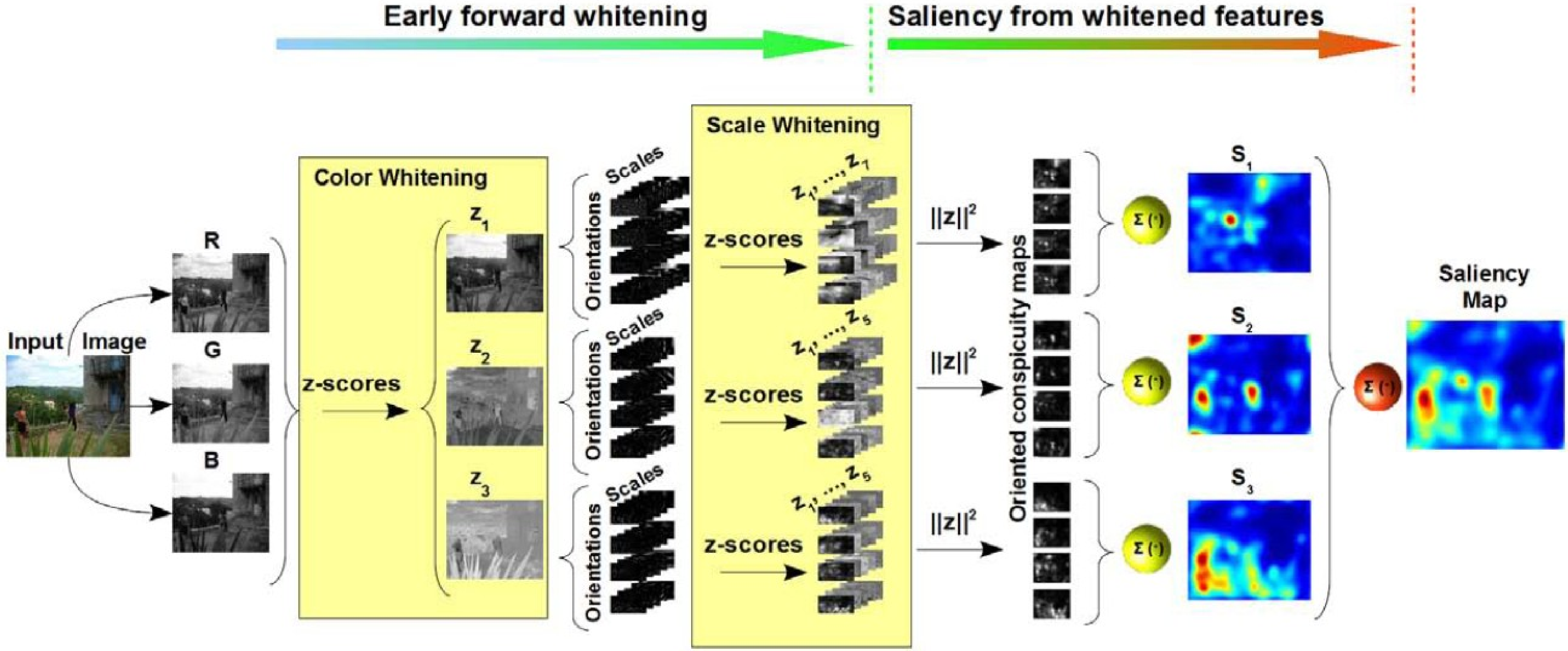
Simone Frintrop

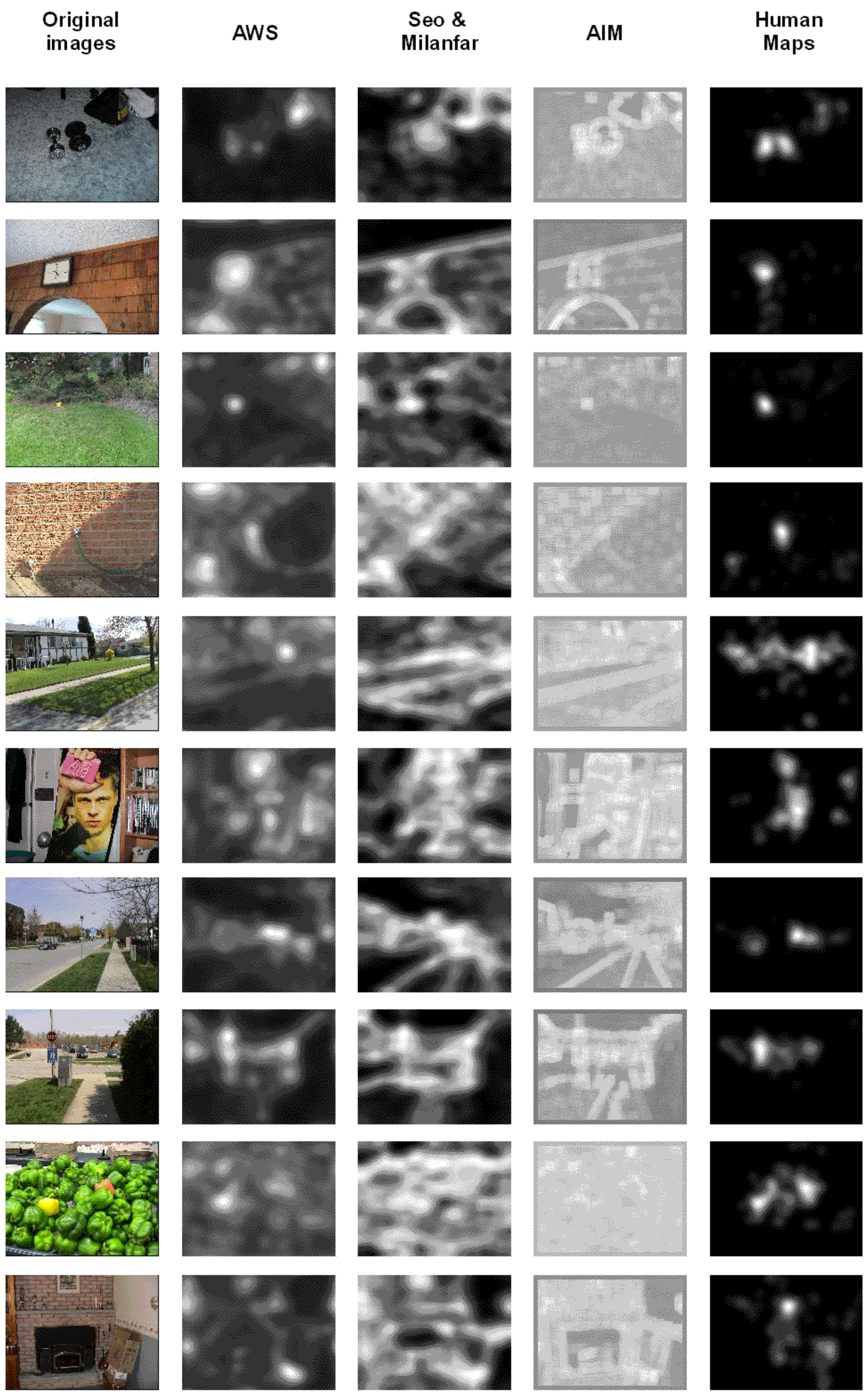




Adaptive Whitening Saliency (AWS)

Garcia Diaz et al., 2011



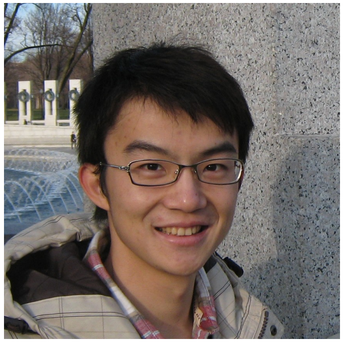


Model	Bruce & Tsotsos Dataset
AWS	0.7106
Seo & Milanfar (JoV 2009)	0.6896**
AIM (JoV 2009)	0.6727*
SUN (JoV 2008)	0.6682*
Itti et al. (PAMI 1998)	0.6456
Gao et al. (2008)	0.6395*

AWS

5 lines of MATLAB codes to implement Spectral Residual

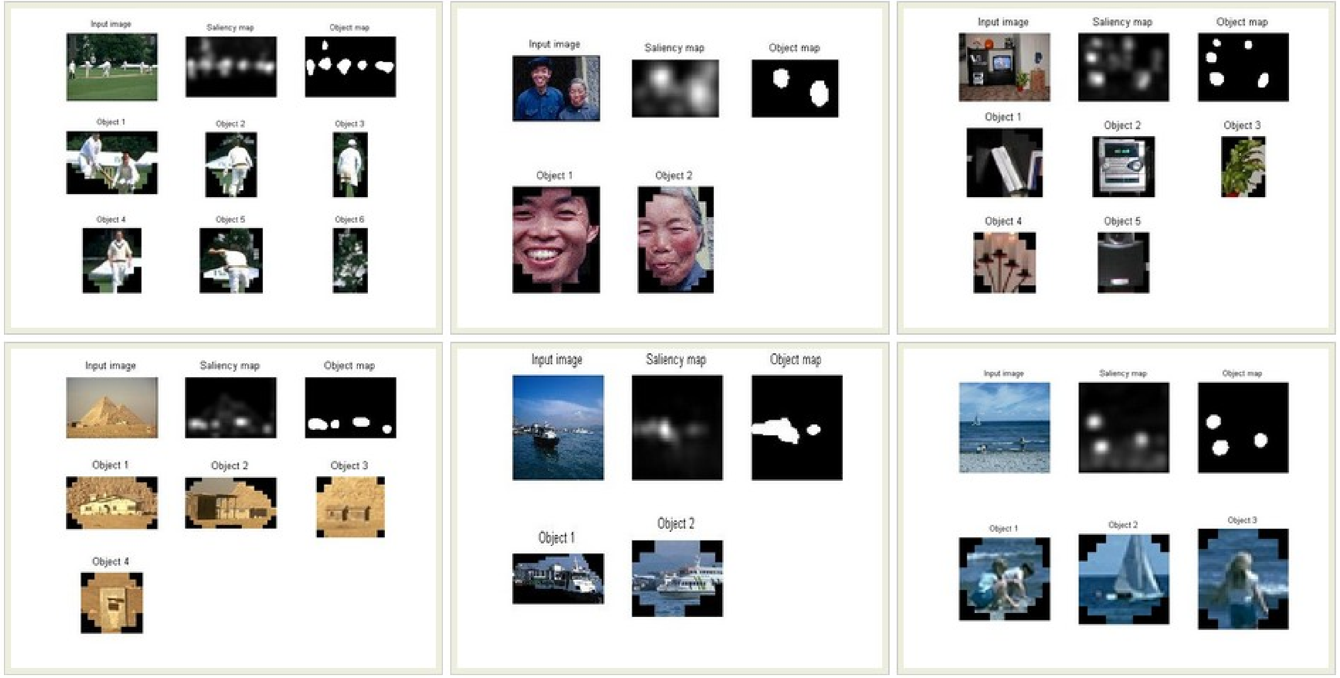
```
clear
clc

%% Read image from file
inImg = im2double(rgb2gray(imread('yourImage.jpg')));
inImg = imresize(inImg, 64/size(inImg, 2));

%% Spectral Residual
myFFT = fft2(inImg);
myLogAmplitude = log(abs(myFFT));
myPhase = angle(myFFT);
mySpectralResidual = myLogAmplitude - imfilter(myLogAmplitude, fspecial('average', 3), 'replicate');
saliencyMap = abs(ifft2(exp(mySpectralResidual + i*myPhase))).^2;

%% After Effect
saliencyMap = mat2gray(imfilter(saliencyMap, fspecial('gaussian', [10, 10], 2.5)));
imshow(saliencyMap);
```

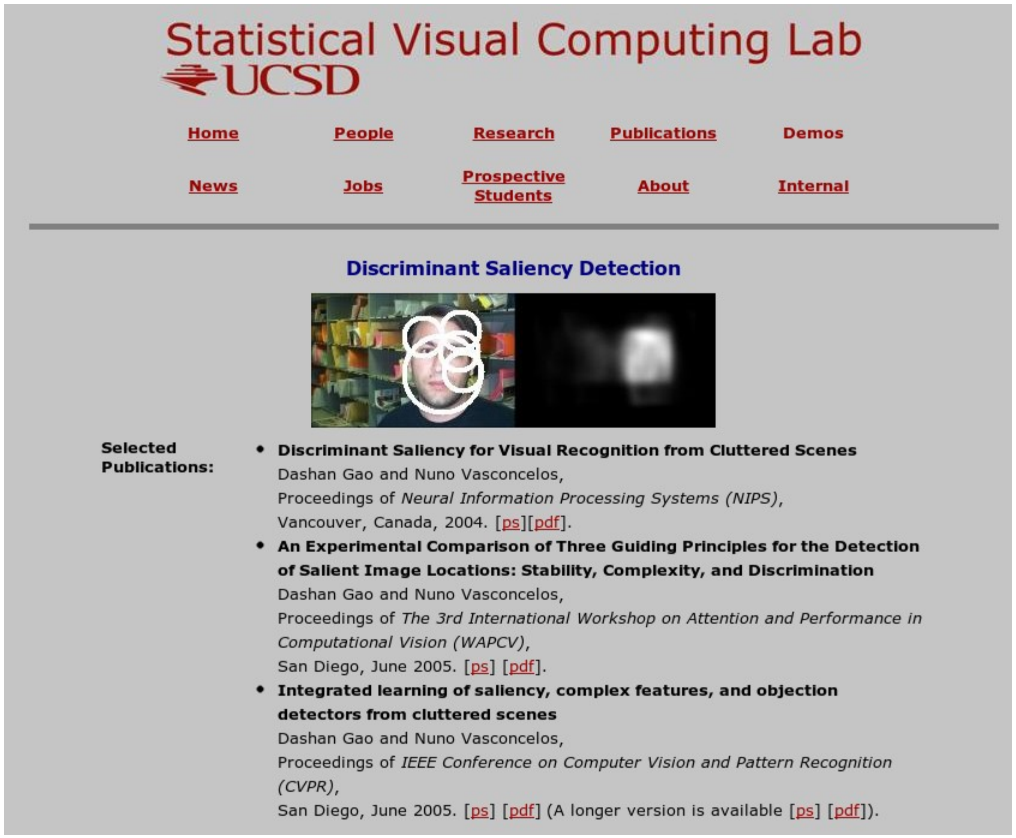


Xiaodi Hou

$$H(\text{Image}) = H(\text{Innovation}) + H(\text{Prior Knowledge})$$

Results on natural images





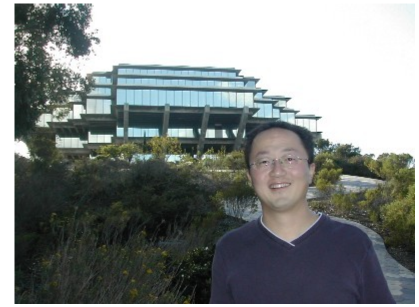
Statistical Visual Computing Lab
UCSD

Home | People | Research | Publications | Demos
News | Jobs | Prospective Students | About | Internal

Discriminant Saliency Detection

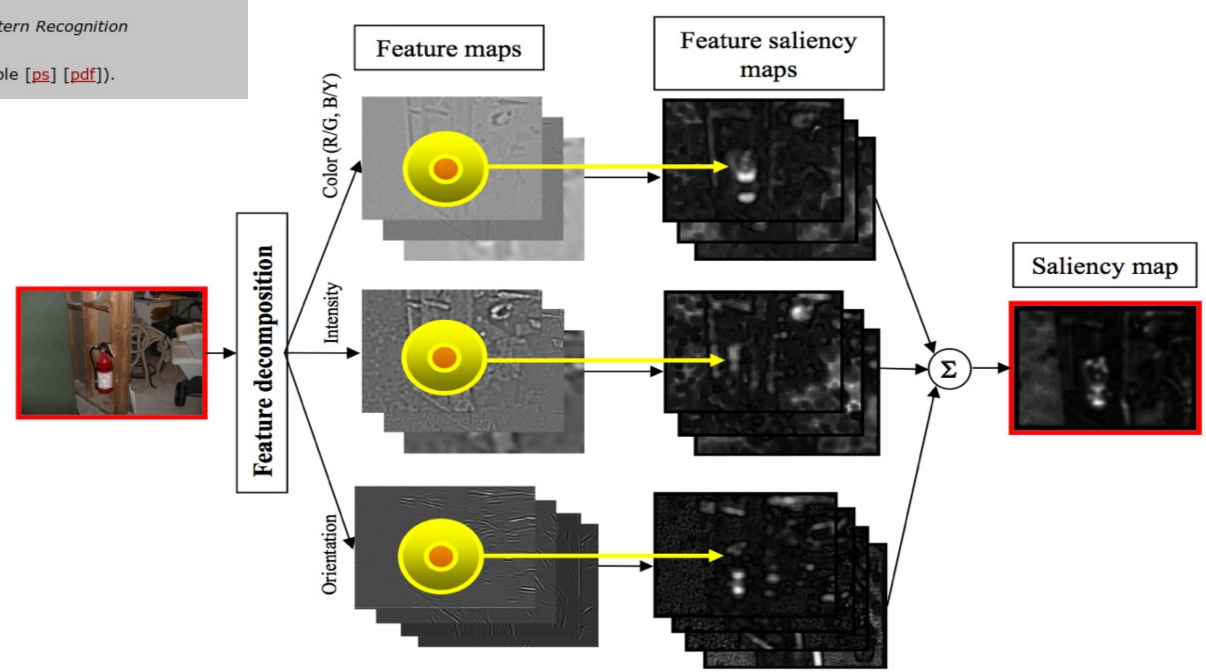
Selected Publications:

- **Discriminant Saliency for Visual Recognition from Cluttered Scenes**
  Dashan Gao and Nuno Vasconcelos,
  Proceedings of *Neural Information Processing Systems (NIPS)*,
  Vancouver, Canada, 2004. [ps][pdf].
- **An Experimental Comparison of Three Guiding Principles for the Detection of Salient Image Locations: Stability, Complexity, and Discrimination**
  Dashan Gao and Nuno Vasconcelos,
  Proceedings of *The 3rd International Workshop on Attention and Performance in Computational Vision (WAPCV)*,
  San Diego, June 2005. [ps] [pdf].
- **Integrated learning of saliency, complex features, and objection detectors from cluttered scenes**
  Dashan Gao and Nuno Vasconcelos,
  Proceedings of *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*,
  San Diego, June 2005. [ps] [pdf] (A longer version is available [ps] [pdf]).



Dashan Gao

Discriminant Saliency Detection




journal of VISION
A Journal of scientific rese
on biological v

Home | About | Current Issue | Back Issues

Institution: Norris Medical Lib

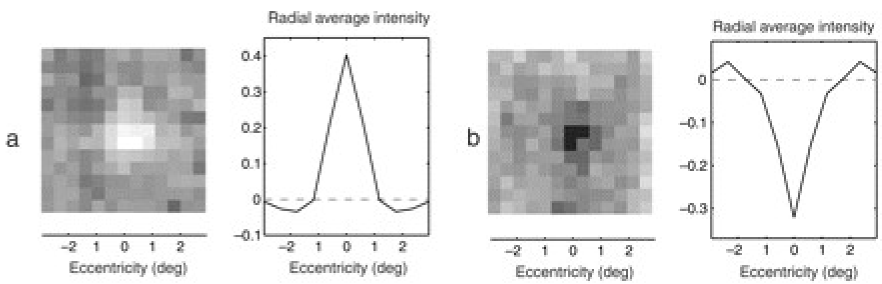
Center-surround patterns emerge as optimal predictors for human saccade targets

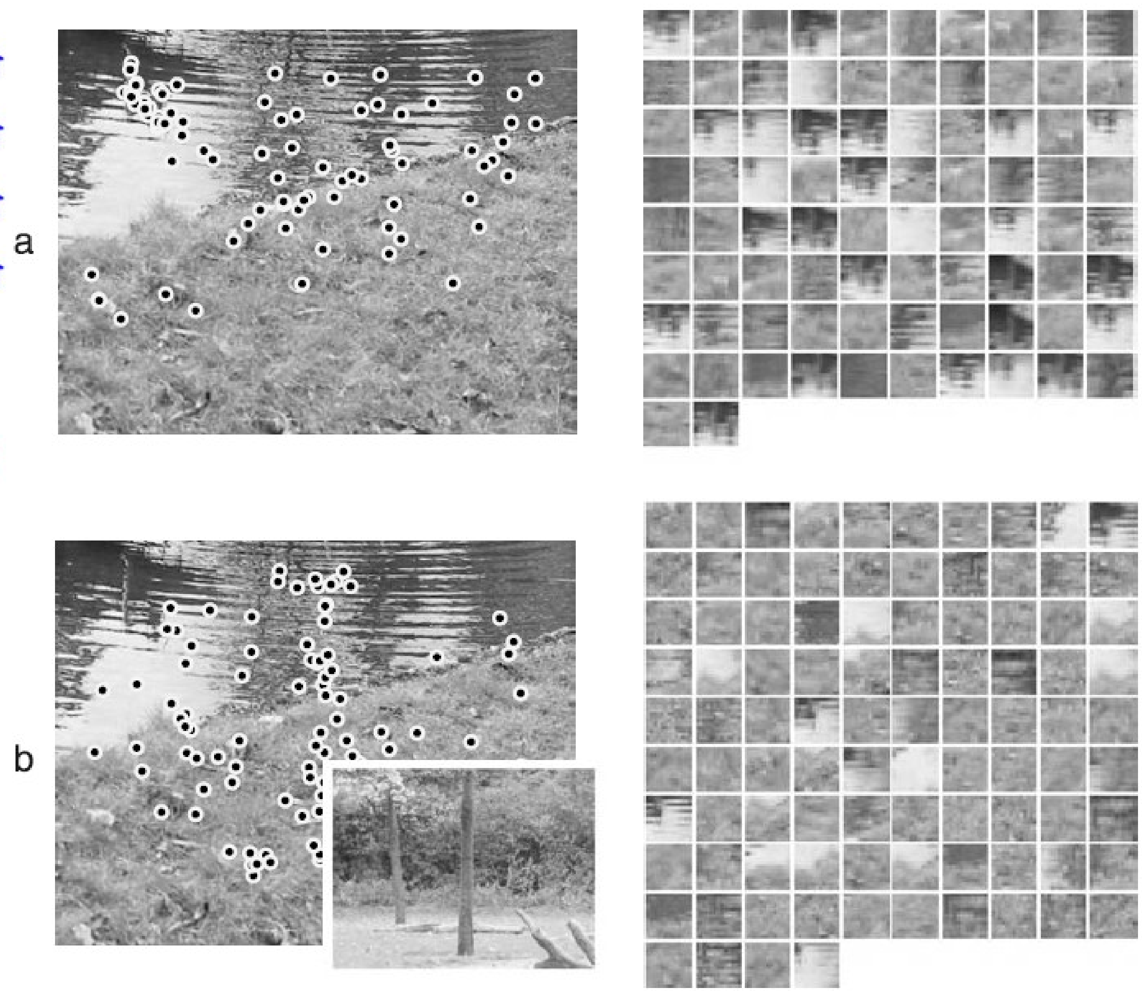
Wolf Kienzle [1],
Matthias O. Franz [2,3],
Bernhard Schölkopf [4] and
Felix A. Wichmann [5,6,7]

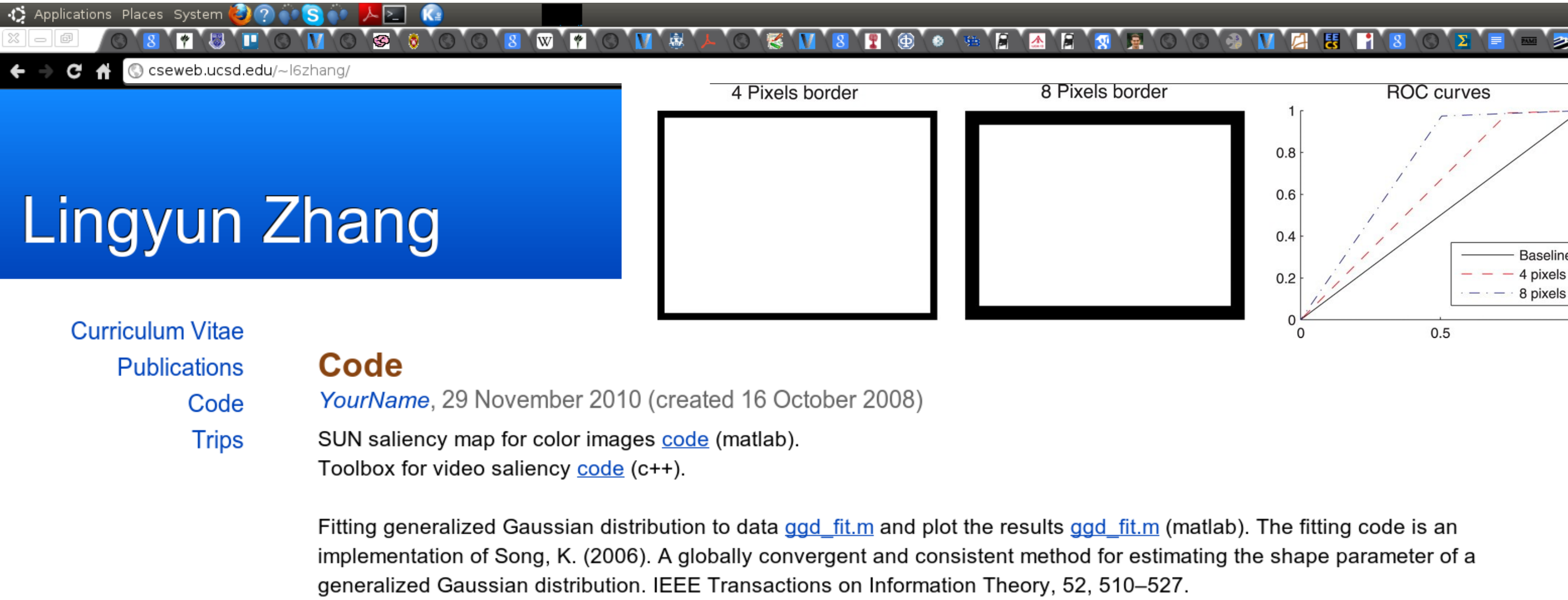
+ Author Affiliations

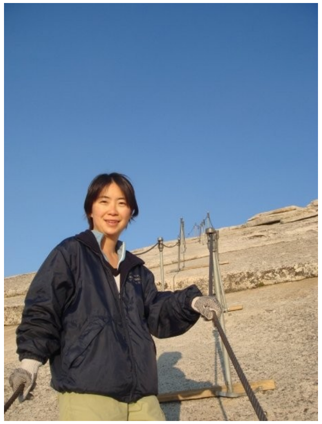
Abstract

Wolfe Kienzle









Lingyun Zhang

SUN model

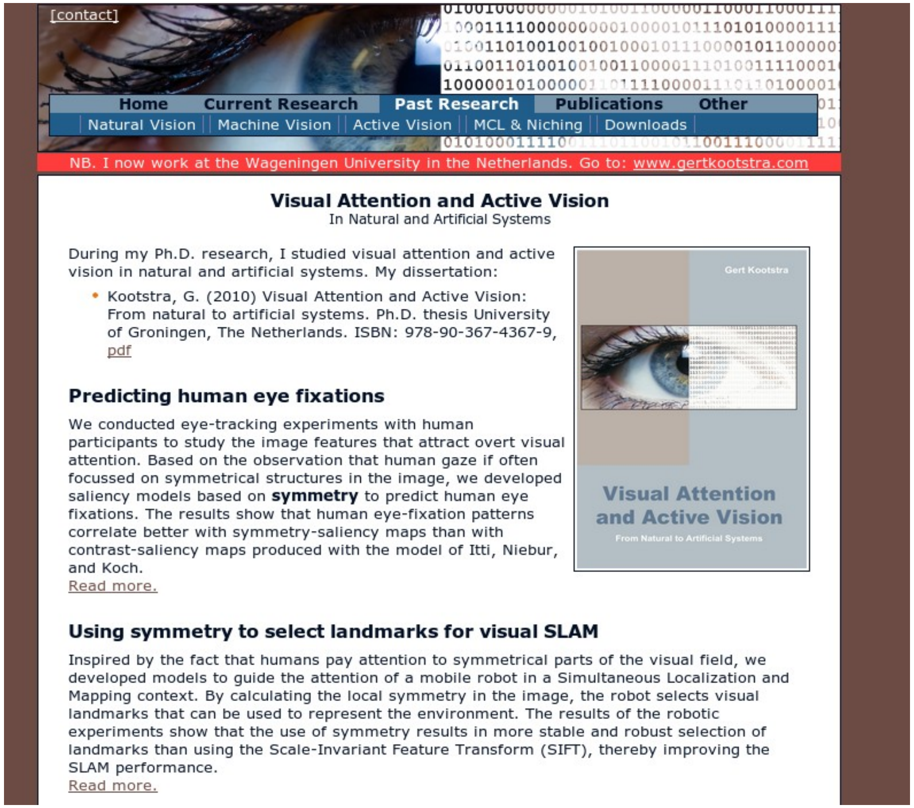
$$\log p(T_x|F_x) = \log \frac{p(F_x|T_x)p(T_x)}{p(F_x)}$$

$$= \log p(F_x|T_x) + \log p(T_x) + \log \frac{1}{p(F_x)}$$

BU Saliency

Paying attention to symmetry

Gert Kootstra



[contact]

Home | Current Research | Past Research | Publications | Other

Natural Vision | Machine Vision | Active Vision | MCL & Niching | Downloads

NB. I now work at the Wageningen University in the Netherlands. Go to: www.gertkootstra.com

Visual Attention and Active Vision

In Natural and Artificial Systems

During my Ph.D. research, I studied visual attention and active vision in natural and artificial systems. My dissertation:

- Kootstra, G. (2010) Visual Attention and Active Vision: From natural to artificial systems. Ph.D. thesis University of Groningen, The Netherlands. ISBN: 978-90-367-4367-9, pdf

Gert Kootstra

Visual Attention and Active Vision

From Natural to Artificial Systems

Predicting human eye fixations

We conducted eye-tracking experiments with human participants to study the image features that attract overt visual attention. Based on the observation that human gaze if often focussed on symmetrical structures in the image, we developed saliency models based on **symmetry** to predict human eye fixations. The results show that human eye-fixation patterns correlate better with symmetry-saliency maps than with contrast-saliency maps produced with the model of Itti, Niebur, and Koch.

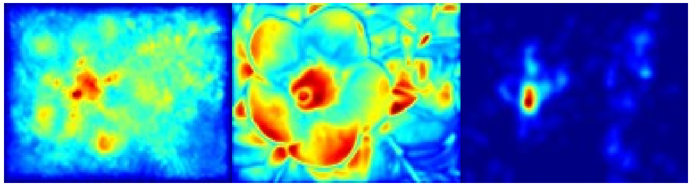
Read more.

Using symmetry to select landmarks for visual SLAM

Inspired by the fact that humans pay attention to symmetrical parts of the visual field, we developed models to guide the attention of a mobile robot in a Simultaneous Localization and Mapping context. By calculating the local symmetry in the image, the robot selects visual landmarks that can be used to represent the environment. The results of the robotic experiments show that the use of symmetry results in more stable and robust selection of landmarks than using the Scale-Invariant Feature Transform (SIFT), thereby improving the SLAM performance.

Read more.

Image | Contrast model | Symmetry model | Human fixations

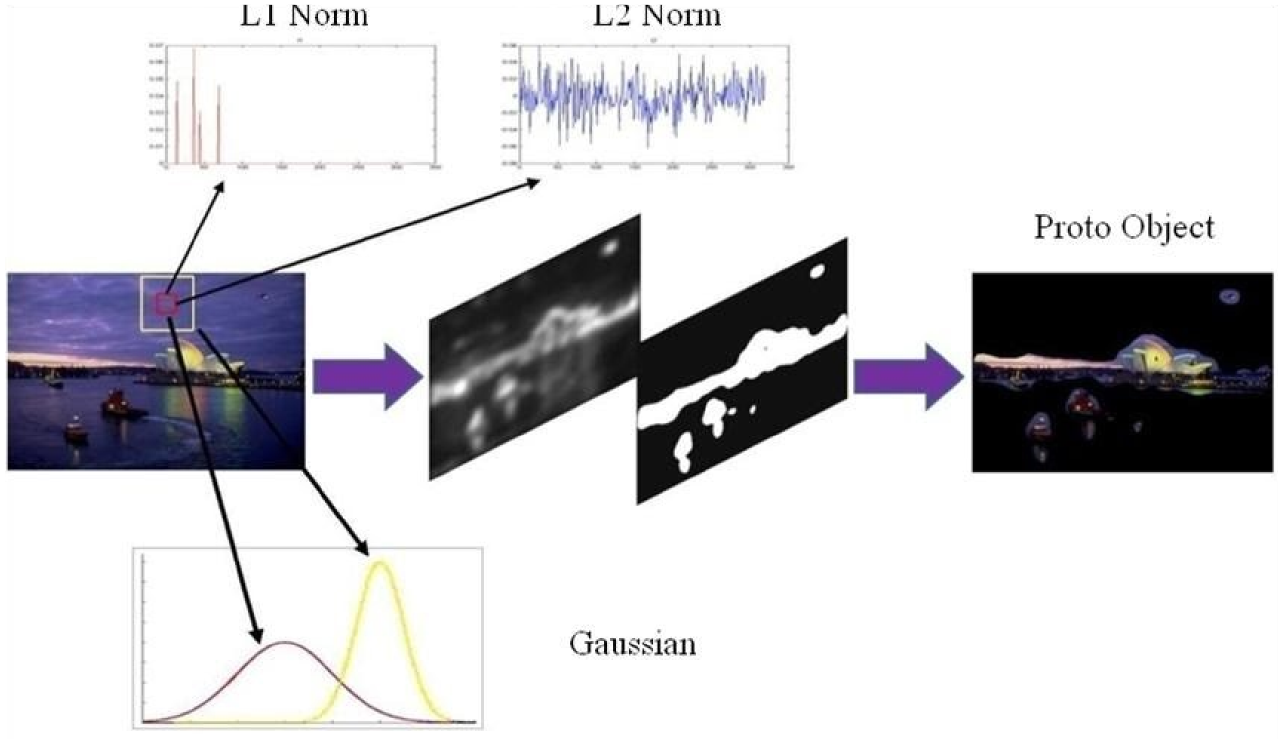


Kootstra | Itti | Human IO

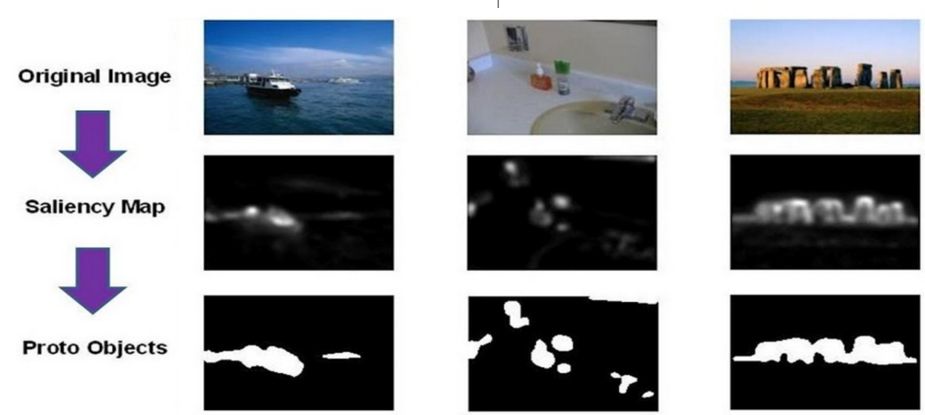
E-Saliency

Tamar Avraham

Visual Saliency Based on Lossy Coding



Yin Li



Yin Li, Yue Zhou, Lei Xu and Xiaochao Yang.

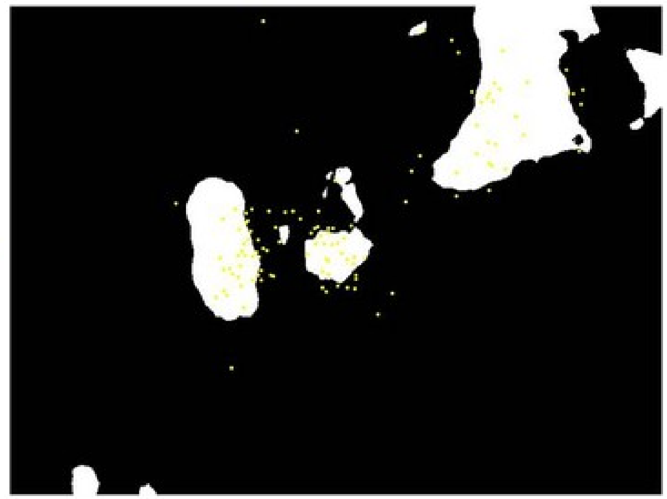
Incremental Sparse Saliency Detection, ICIP 2009

Bottom-Up Visual Attention

In this project, we obtain saliency maps from color images using perceptual characteristics of the CIWaM (Chromatic Induction Wavelet Model).

Example

Original image with eye fixations in yellow



Saliency map of image showing 10% most salient regions

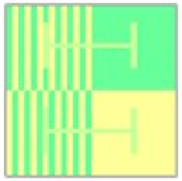
Naila Murray

Matlab code for saliency estimation

Matlab code for our saliency estimation method can be found **here**. For a quick start run "SIM_demo.m." Information can be found in the README file.

References

Naila Murray, Maria Vanrell, Xavier Otazu and C. Alejandro Párraga. **Saliency Estimation Using a Non-Parametric Low-Level Vision Model.** *(to appear) CVPR 2011.*



Toward a unified chromatic induction model

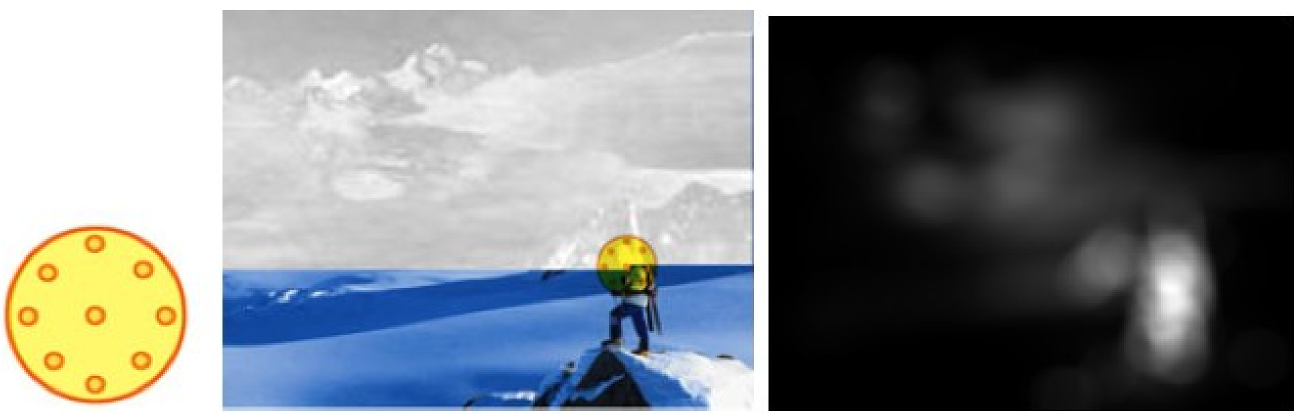
X. Otazu, C.A. Parraga, & M. Vanrell

Journal of Vision 2010, Volume 10, Number 12, Article 5

Fast and Efficient Saliency Detection

(a) (b)

Hamed R. Tavakkoli



(a) (b) (c)

Rezazadegan Tavakoli H, Rahtu E & Heikkilä J ,
Fast and Efficient Saliency Detection Using Sparse Sampling and Kernel Density Estimation,
Image Analysis, SCIA 2011

What and where: A Bayesian inference theory of visual attention

Sharat Chikkerur

S. Chikkerur, T. Serre, C. Tan and Poggio, "What and Where: A Bayesian inference theory of visual attention", Vision Research, 2010

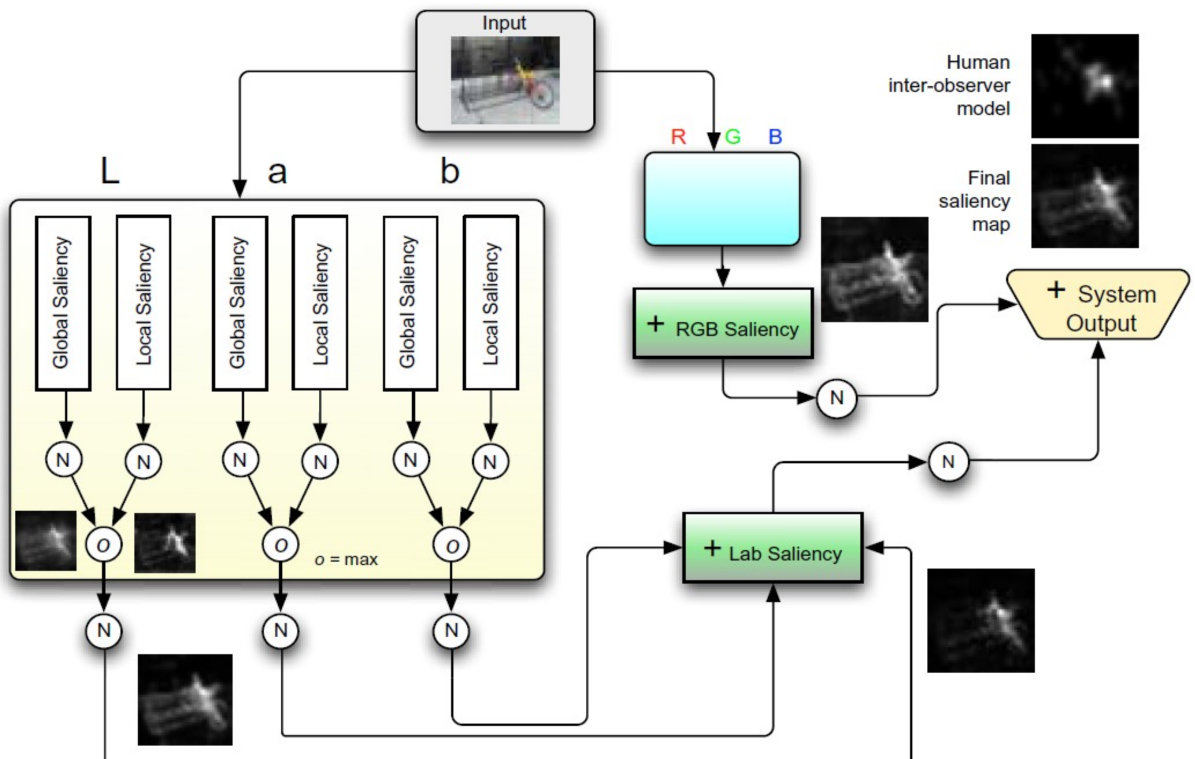
Exploiting Local and Global Patch Rarities for Saliency Detection

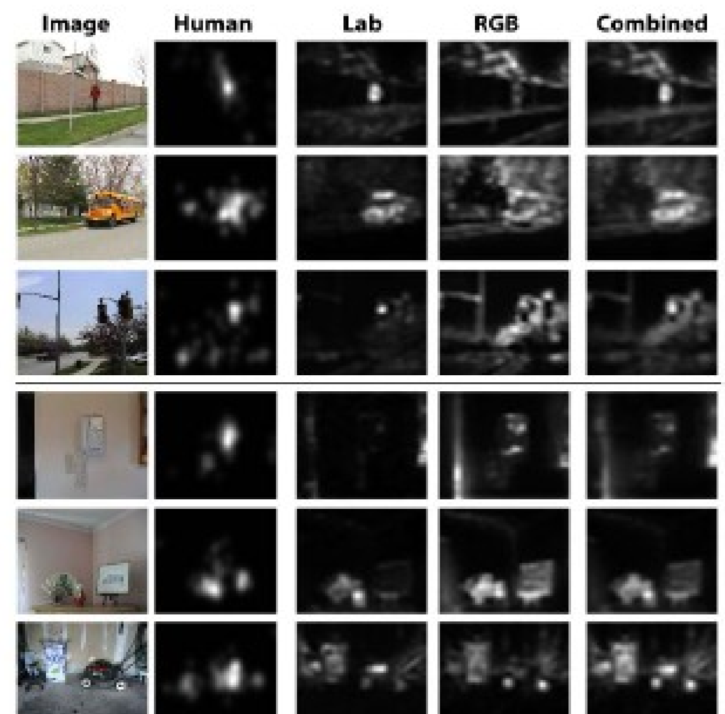
Ali Borji
USC
borji@usc.edu

Laurent Itti
USC
itti@usc.edu

CVPR 2012

Ali Borji

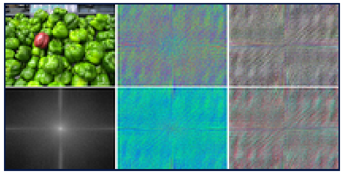




Quaternion-based Spectral Saliency

Boris Schauerte

B. Schauerte, R. Stiefelhagen, "***Quaternion-based Spectral Saliency Detection for Eye Fixation Prediction***". In Proceedings of the 12th European Conference on Computer Vision (ECCV), Firenze, Italy, October 7-13, 2012.



Abstract: In recent years, several authors have reported that spectral saliency detection methods provide state-of-the-art performance in predicting human gaze in images. We systematically integrate and evaluate quaternion DCT- and FFT-based spectral saliency detection, weighted quaternion color space components, and the use of multiple resolutions. Furthermore, we propose the use of the eigenaxes and eigenangles for spectral saliency models that are based on the quaternion Fourier transform. We demonstrate the outstanding performance on the Bruce-Tsotsos (Toronto), Judd (MIT), and Kootstra-Schomacker eye-tracking data sets.

Keywords: Spectral Saliency, Quaternion; Multi-Scale, Color Space, Quaternion Component Weight, Quaternion Axis; Attention; Human Gaze, Eye-Tracking; Bruce-Tsotsos (Toronto), Judd (MIT), and Kootstra-Schomacker data set

Download: [pdf] [bibtex] [code #1 - visual saliency toolbox] [code #2 - Matlab AUC measure implementation] [poster (1)] [poster (2)]

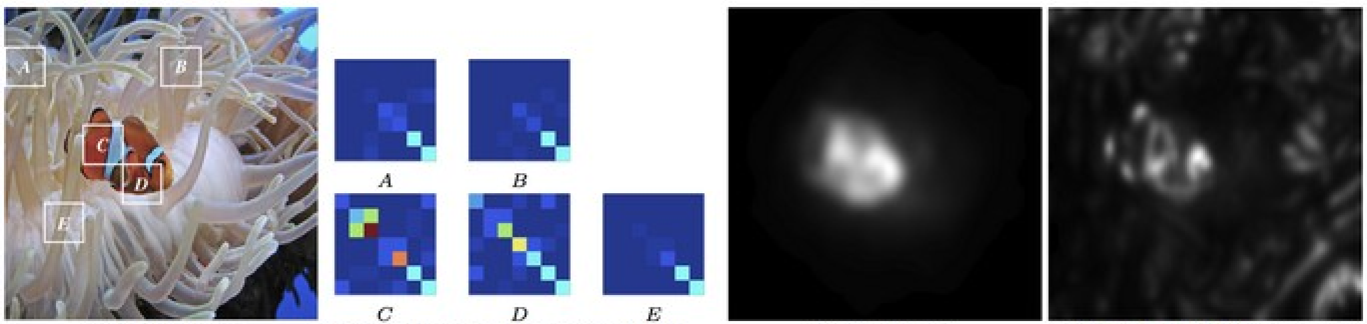
Selected Related Publications:

- Predicting Human Gaze using Quaternion DCT Image Signature Saliency and Face Detection, 2012

Erkut Erdem

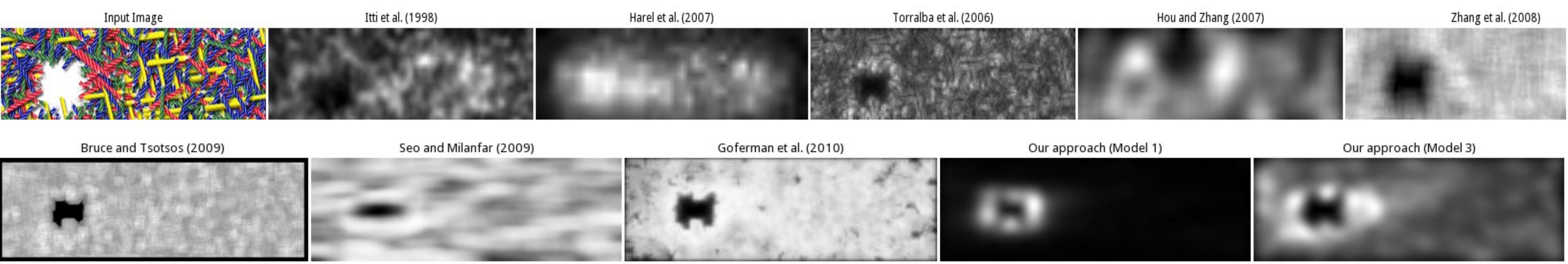
Visual saliency estimation by nonlinearly integrating features using region covariances

Erkut Erdem, Aykut Erdem

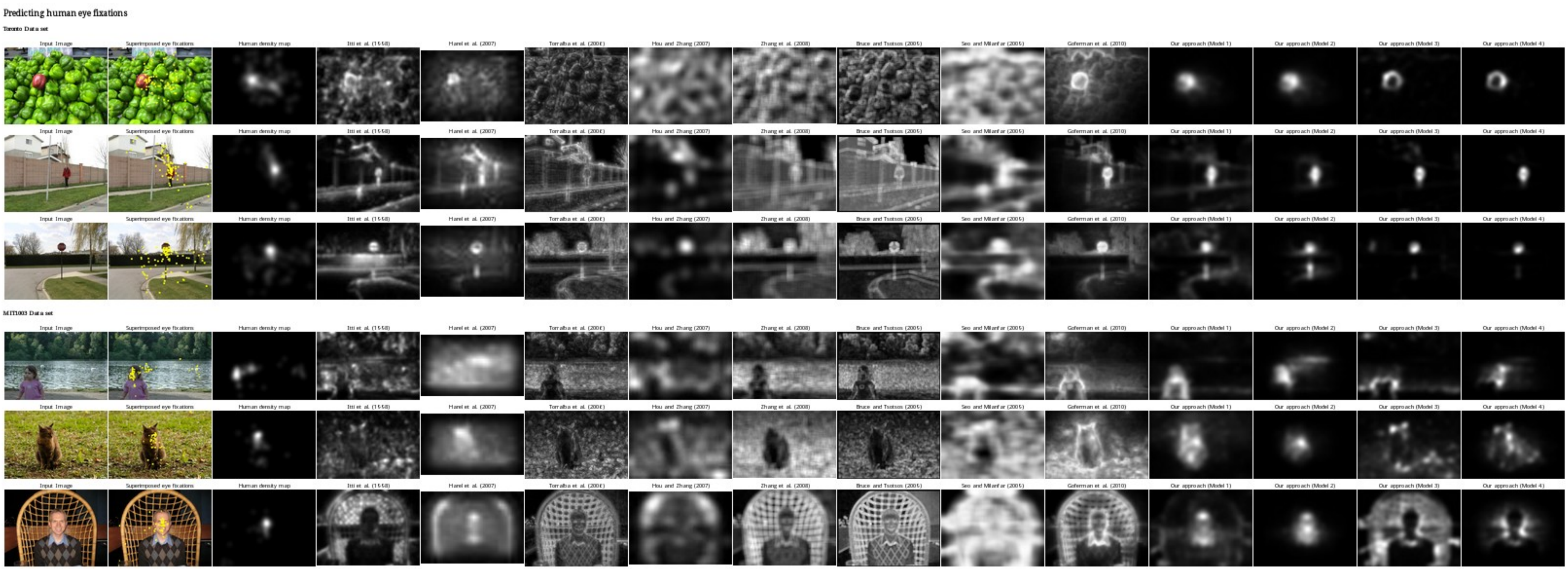


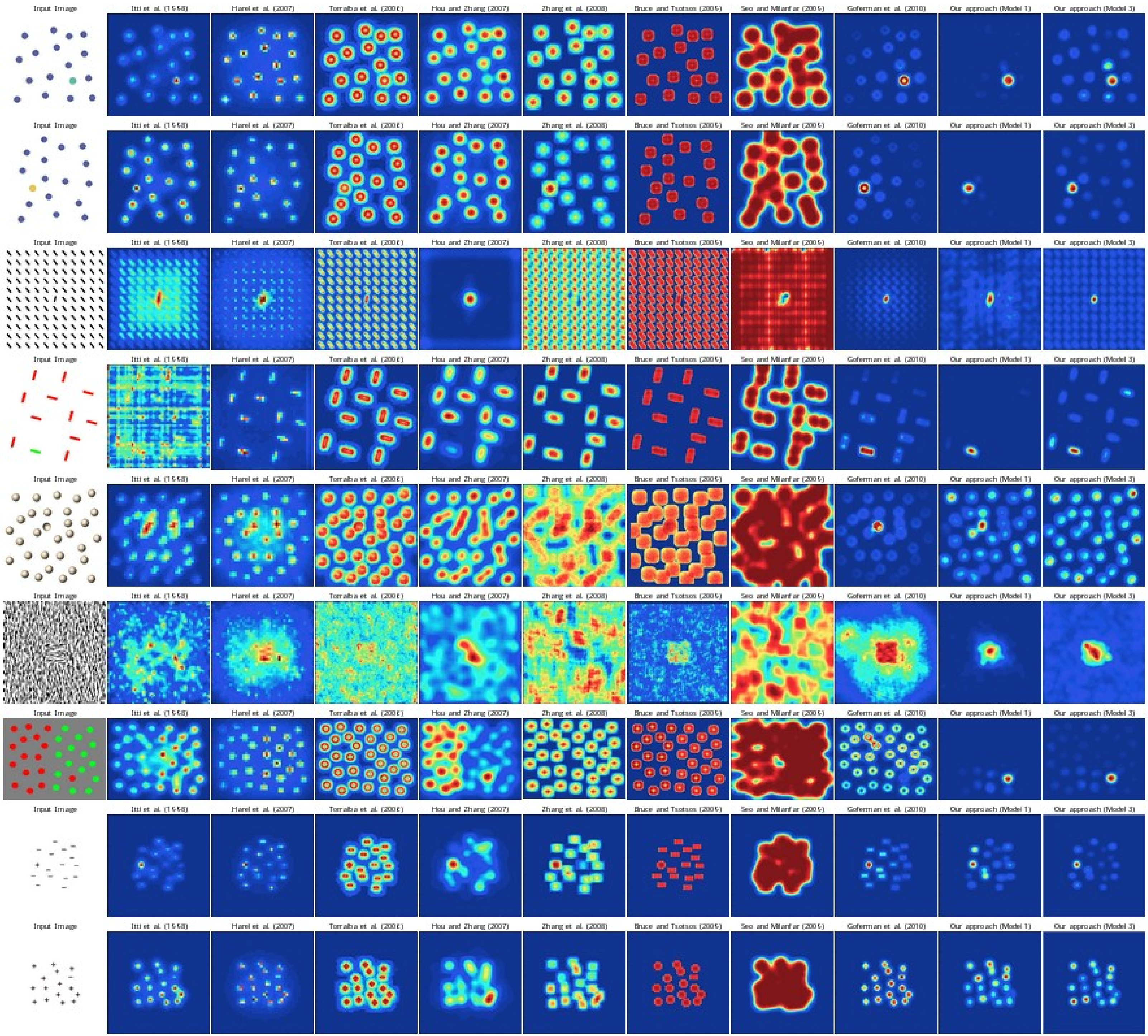
(a) Input image (b) Covariance matrices for the regions A-E (c) Our result (d) Itti et al. (1998)'s result

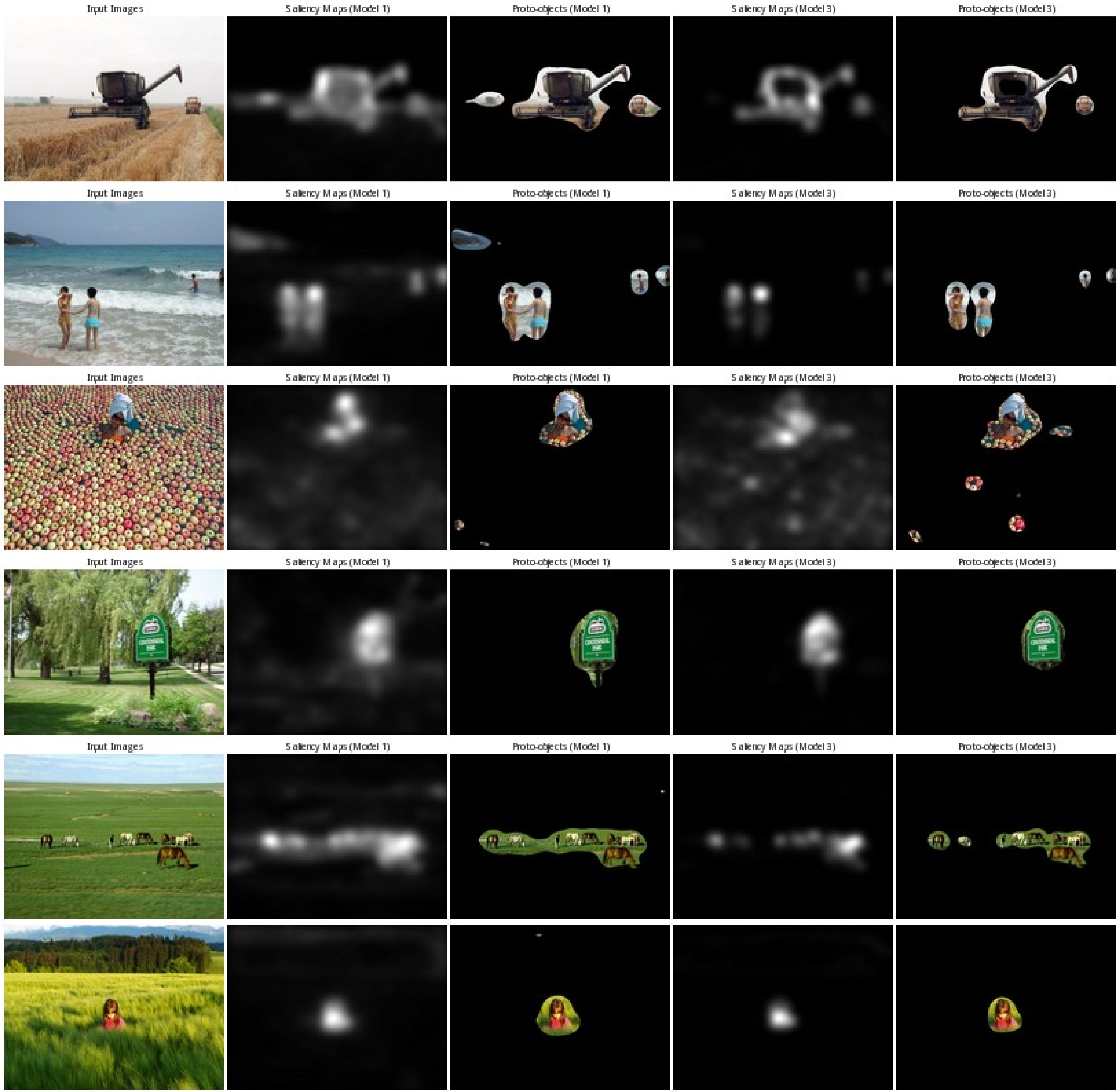
The proposed saliency model. The input image is first decomposed into non-overlapping regions, and then the saliency of each region is measured by examining its surrounding regions. The salient regions are those that are highly dissimilar to their neighboring regions in terms of their covariance representations based on color, orientation, and spatial features. In the saliency map computed by the proposed model, the fish pops out from the complex background, cf. Itti's saliency map (Itti, Koch, & Niebur, 1998).



Erdem and Erdem, JOV 2013




Input Image
Harel et al. (2007)
Hou and Zhang (2007)
Zhang et al. (2008)
Goferman et al. (2010)
Our approach (Model 1)
Our approach (Model 3)


Input Images
Saliency Maps (Model 1)
Proto-objects (Model 1)
Saliency Maps (Model 3)
Proto-objects (Model 3)

UNIVERSITY OF CALIFORNIA,
SANTA CRUZ

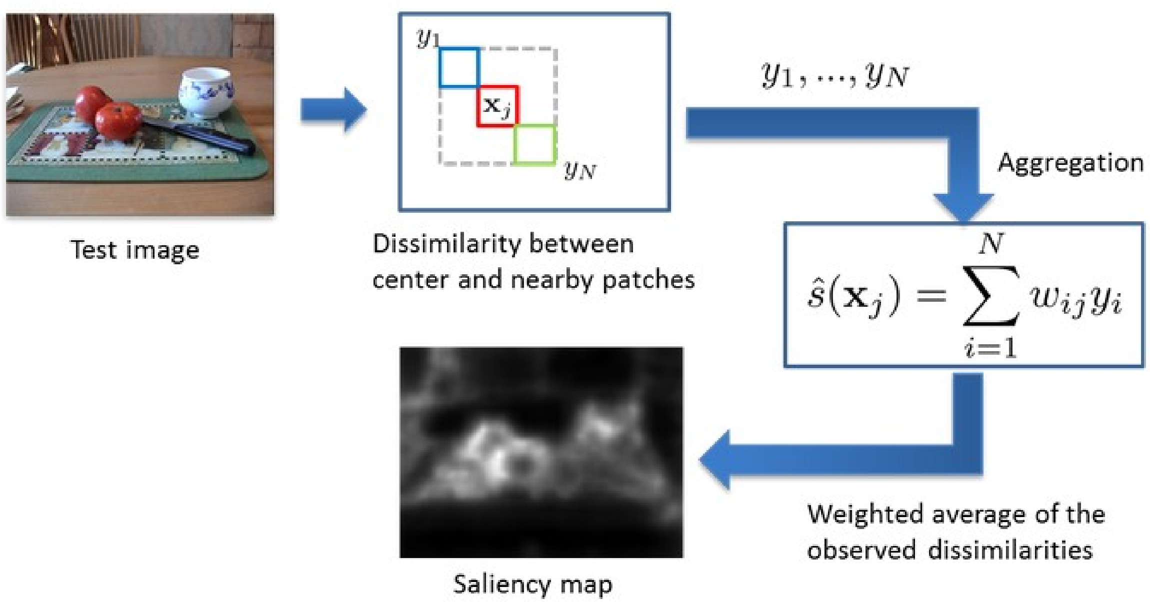
Chelhwon Kim

Visual Saliency in Noisy Images

Chelhwon Kim and **Peyman Milanfar**

- Chelhwon Kim and Peyman Milanfar, **"Visual Saliency in Noisy Images"** ***Journal of Vision*** **13(4):5, March 11, 2013.**
- Chelhwon Kim, and Peyman Milanfar, **"Finding Saliency in Noisy Images"**, ***SPIE Conference on Computational Imaging (8269) , January 2012, Burlingame, CA***

Overview of saliency detection



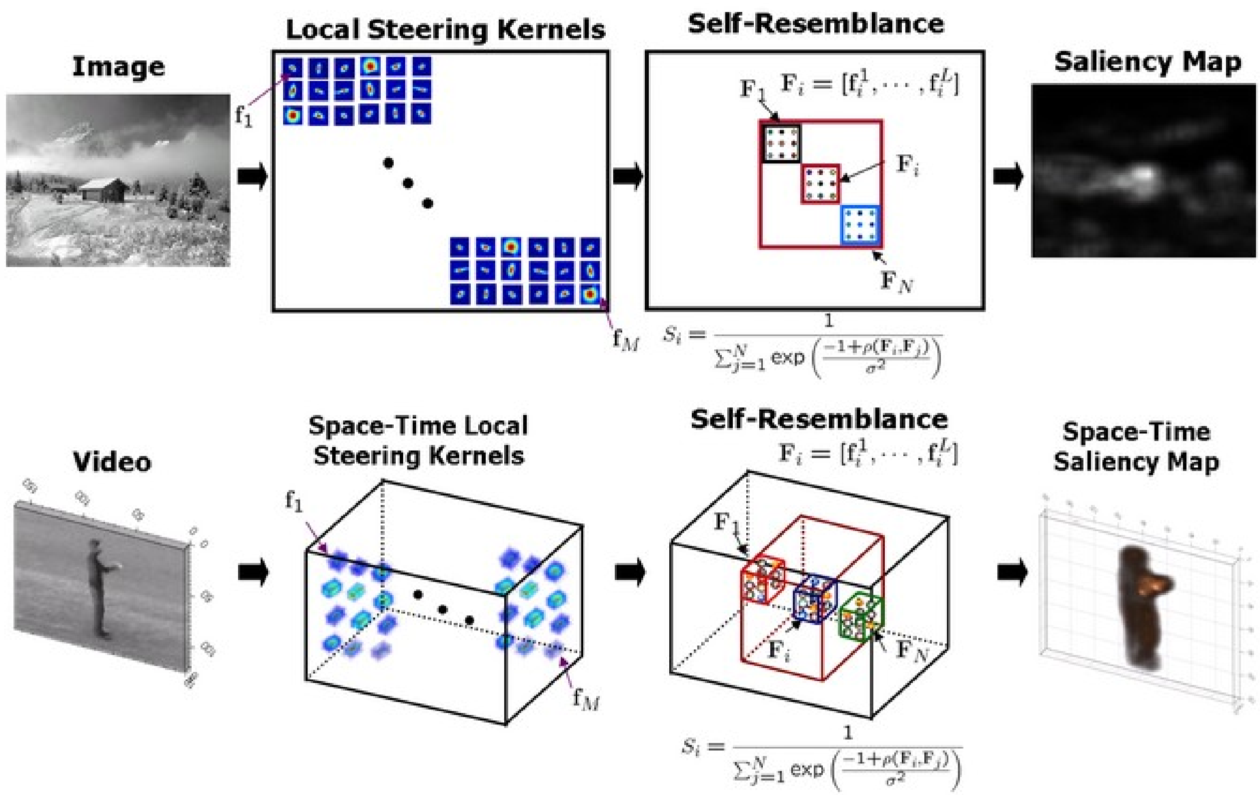
UNIVERSITY OF CALIFORNIA,
SANTA CRUZ

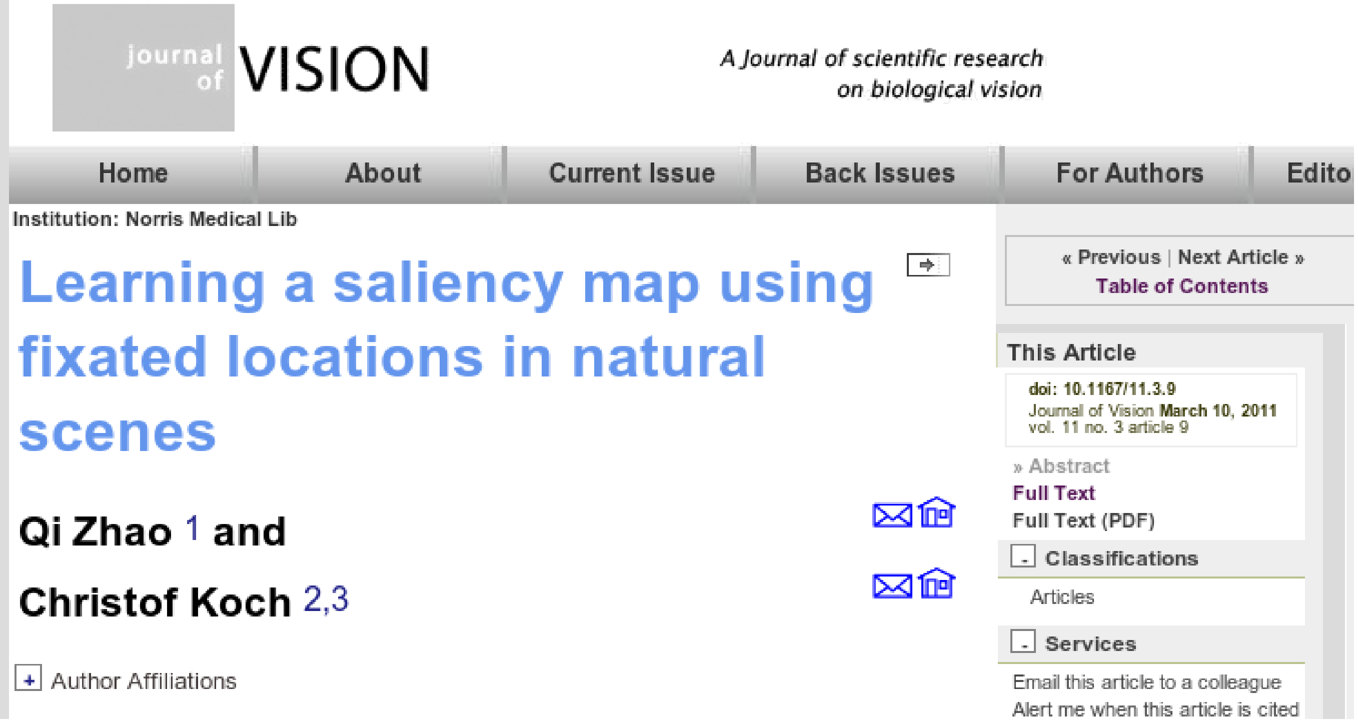
Static and Space-time Visual Saliency Detection by Self-Resemblance

Hae Jong Seo and Peyman Milanfar

- Hae Jong Seo, and Peyman Milanfar, [" Nonparametric Bottom-Up Saliency Detection by Self-Resemblance"], *Accepted for IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 1st International Workshop on Visual Scene Understanding(ViSU), Miami, June,2009*
- Hae Jong Seo, and Peyman Milanfar, ["Static and Space-time Visual Saliency Detection by Self-Resemblance"], *The Journal of Vision 9(12):15, 1-27, http://journalofvision.org/9/12/15/, doi:10.1167/9.12.15*

Hae Jong




journal of VISION

A Journal of scientific research on biological vision

Home | About | Current Issue | Back Issues | For Authors | Edito

Institution: Norris Medical Lib

Learning a saliency map using fixated locations in natural scenes

Qi Zhao 1 **and**

Christof Koch 2,3

Author Affiliations

« Previous | Next Article »
Table of Contents

This Article

doi: 10.1167/11.3.9
Journal of Vision March 10, 2011
vol. 11 no. 3 article 9

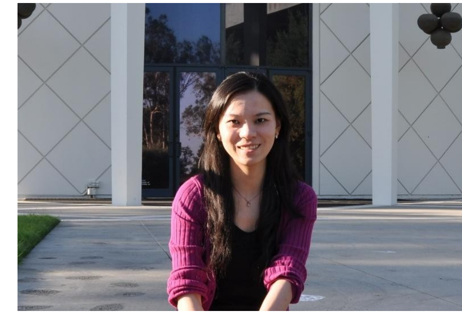
» Abstract
Full Text
Full Text (PDF)

Classifications

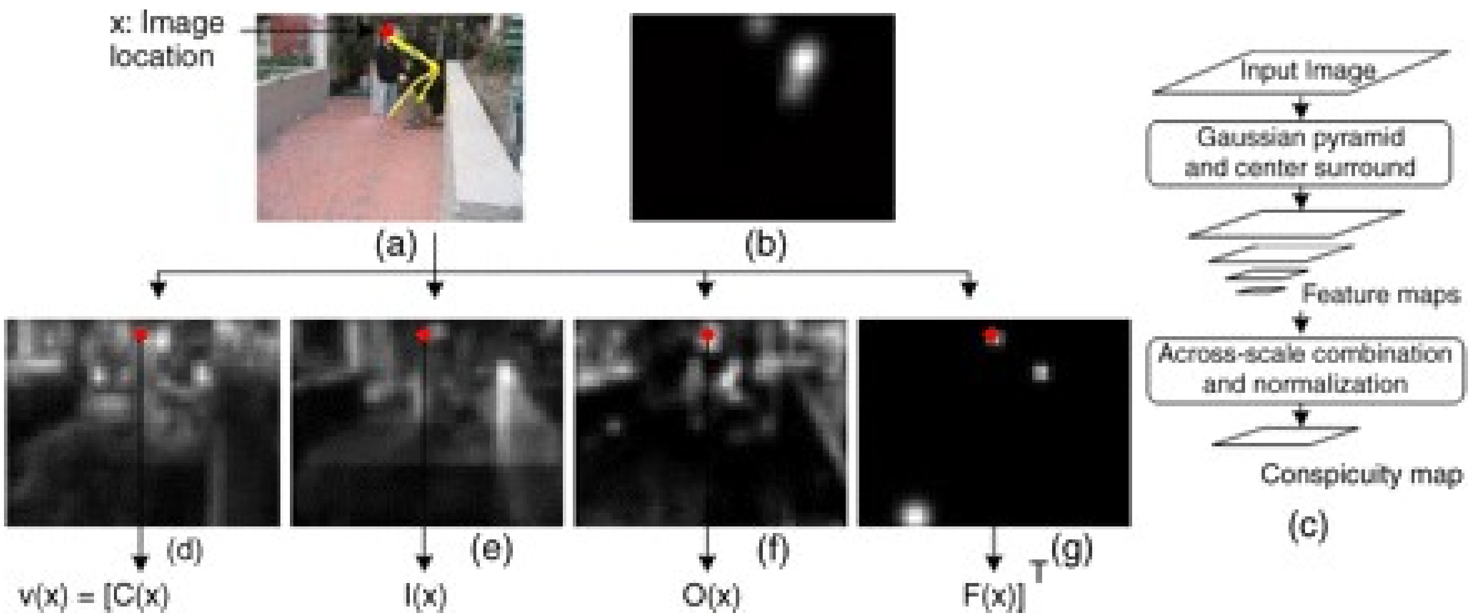
Articles

Services

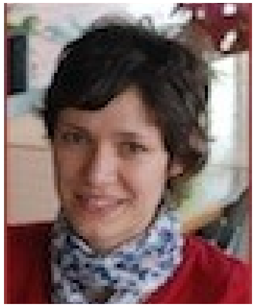
Email this article to a colleague
Alert me when this article is cited



Qi zhao



Spatio-temporal (Dynamic Saliency)



Elenora Vig

- Behavioral studies
  - Variability of eye movements when viewing dynamic natural scenes (Dorr et al., JOV 2013)
  - Eye guidance in natural vision: Reinterpreting salience (Tatler et al. JOV 2010)
  - Adaptive Gaze Control in Natural Environments (Jovancevic-Misic and Hayhoe)
  - Human eye–head co-ordination in natural exploration (Einhauser et al., 2007)
  - Combining top-down processes to guide eye movements during real-world scene search (Malcolm and Henderson; JOV 2010)
  - Viewing task influences eye movement control during active scene perception (Castelhano et al., JOV 2009)
  - Eye movements while viewing narrated, captioned, and silent videos (Ros and Kowler; JOV 2013)
  - Temporal eye movement strategies during naturalistic viewing (Wang et al., JOV 2012)
- Models (See Borji et al., TIP 2013)
  - Usually extension of spatial models
  - Itti plus motion and flickers channels (ITTI + FI)
    - See also VOCUS by Simone Frintrop
  - Le Meur et al., VR; 2007
  - Marat et al., IJCV 2009
  - Seo & Milanfar; JOV 2009
  - Guo et al., TIP
  - SUNDAY; Zhang et al. (SUN++)
  - Peters and Itti; CVPR 2007
  - Borji et al.; CVPR 2012
  - Mathe and Sminchisescu; ECCV 2012
  - Vig et al., ECCV 2012, PAMI 2012
  - Rudoy et al., CVPR 2013
  - Riche et al., ACCV 2012 [A Comparative Study on Videos]
  - Kienzle et al., DAGM 2007
  - Jia Li, et al., IJCV 2010
  - Pang et al,, ICME 2008

Action Recognition

Fixation Prediction

Segmentation

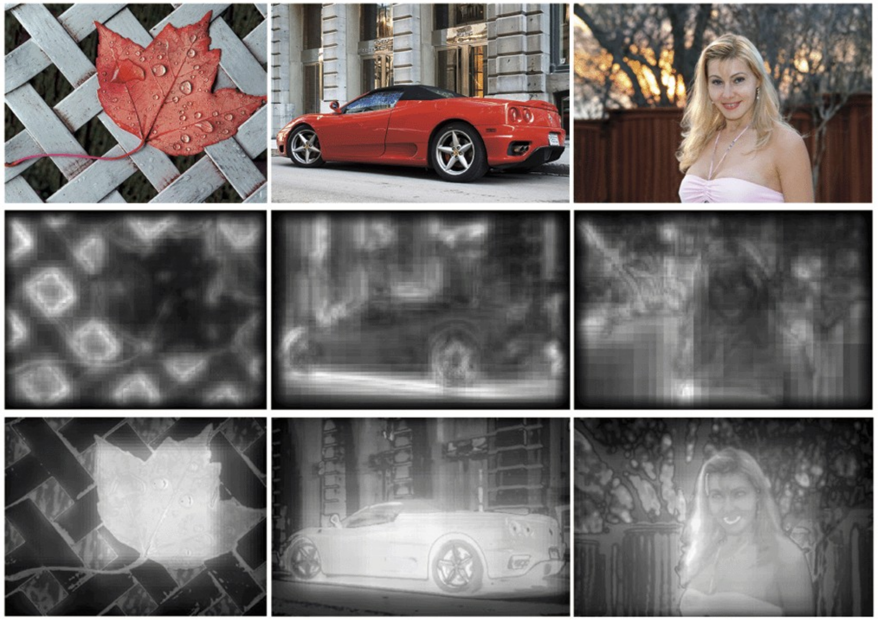
Sophie Marat

Jia Li

Akisato Kimura

Salient Object Detection Models

The story began in 2007:
Learning to Detect a Salient Object
Liu et al., CVPR 2007, PAMI 2011



(a) (b) (c)

Radhakrishna Achanta, Sheila Hemami, Francisco Estrada, and Sabine Susstrunk

Abstract

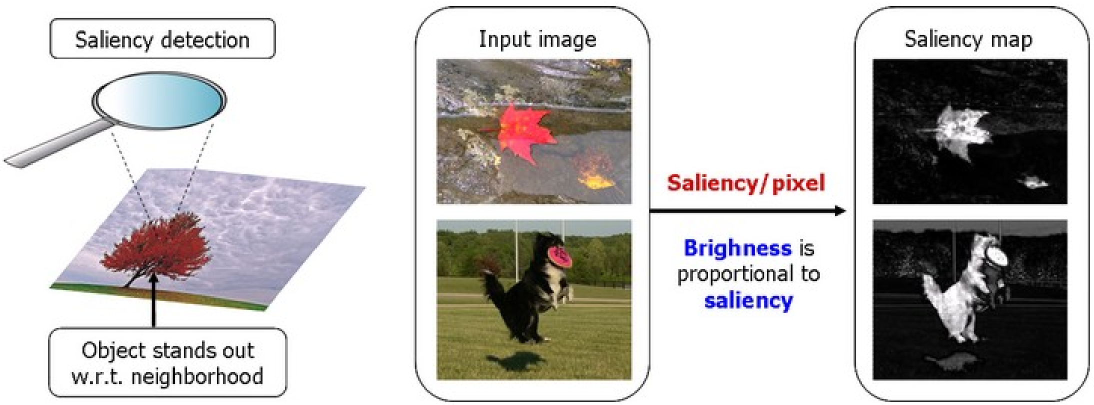
Detection of visually salient image regions is useful for applications like object segmentation, adaptive compression, and object recognition. In this paper, we introduce a method for salient regio with well-defined boundaries of salient objects. These boundaries are preserved by retaining substantially more frequency content from the original image than other existing techniques. Our me simple to implement, and is computationally efficient. We compare our algorithm to five state-of-the-art salient region detection methods with a frequency domain analysis, ground truth, and a s outperforms the five algorithms both on the ground truth evaluation and on the segmentation task by achieving both higher precision and better recall.

Reference and PDF

R. Achanta, S. Hemami, F. Estrada and S. Süsstrunk, **Frequency-tuned Salient Region Detection**, *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2009)*, pp
[**detailed record**] [**bibtex**]

Introduction

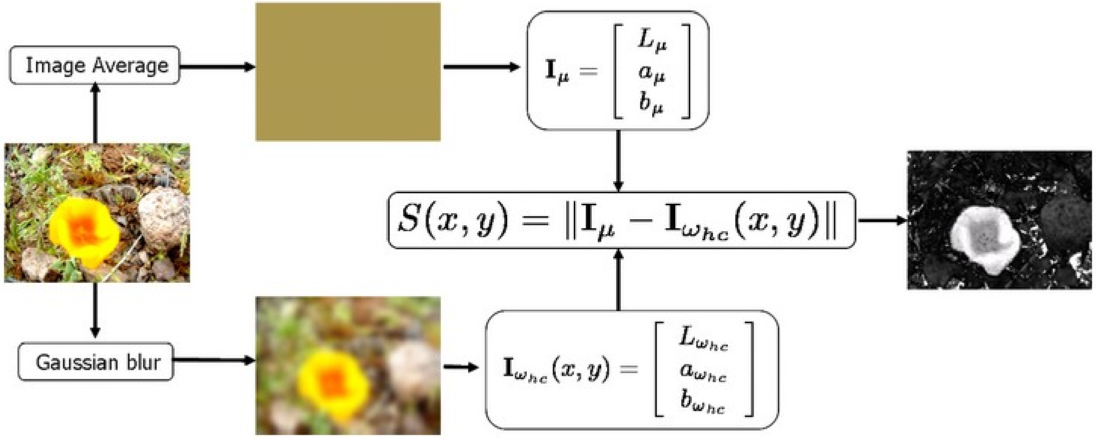
Salient regions and objects *stand-out* with respect to their neighborhood. The goal of our work was to compute the degree of standing out or *saliency* of each pixel with respect to its neighbourhc saliency detection methods take a similar center-versus-surround approach. One of the key decisions to make is the size of the neighborhood used for computing saliency. In our case we use the exploit more spatial frequencies than state-of-the-art methods (please refer to the paper for details) resulting in uniformly highlighted salient regions with well-defined borders.



Radhakrishna Achanta

Saliency Detection Algorithm

In simple words, our method find the Euclidean distance between the *Lab* pixel vector in a Gaussian filtered image with the average *Lab* vector for the input image. This is illustrated in the figure provide a comparison of our method with state-of-the-art methods.



Models	Challenges	Benchmark	Summary

Salient Region Detection and Segmentation

Author

Radhakrishna Achanta

Goal of Research

The goal of this research is to detect and segment salient regions in an image.

Abstract

Detection of salient image regions is useful for applications like image segmentation, adaptive compression, and region-based image retrieval. In this paper we present a novel method to d of luminance and color. The method is fast, easy to implement and generates high quality saliency maps of the same size and resolution as the input image.We demonstrate the use of the a whole objects from digital images.

Publication

R. Achanta, F. Estrada, P. Wils and S. Süsstrunk, **Salient Region Detection and Segmentation**, *International Conference on Computer Vision Systems (ICVS '08)*, Vol. 5008, Springer l
[**detailed record**] [**bibtex**]

Download the Salient Region Detector (for Windows only)

Windows executable (GUI based)
Windows executable (command line)

Download MATLAB code

Saliency_ICVS_2008.m

Results

The results of our work are shown below. The images below, from left to right, are: the original image, saliency map using Itti's method, segmentation result using Itti's map, saliency map using our method.


Stas Gofferman

Global Contrast-based Salient Region Detection

Ming-Ming Cheng[1] Guo-Xin Zhang[1] Niloy J. Mitra[2] Xiaolei Huang[3] Shi-Min Hu[1]

[1]TNList, Tsinghua University, Beijing [2]KAUST/IIT Delhi [3]Lehigh University

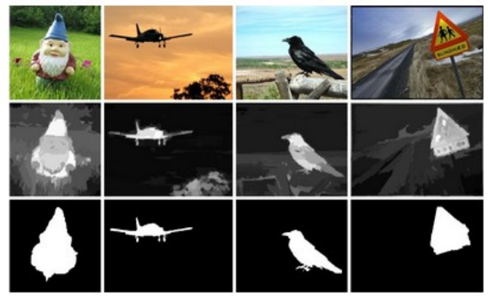


Figure. Given input images (top), a global contrast analysis is used to compute high resolution saliency maps (middle), which can be used to produce masks (bottom) around regions of interest.

Ming-Ming Cheng

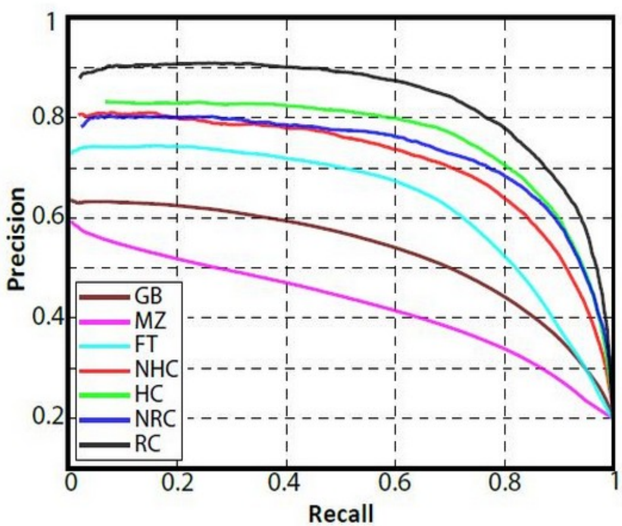
Abstract

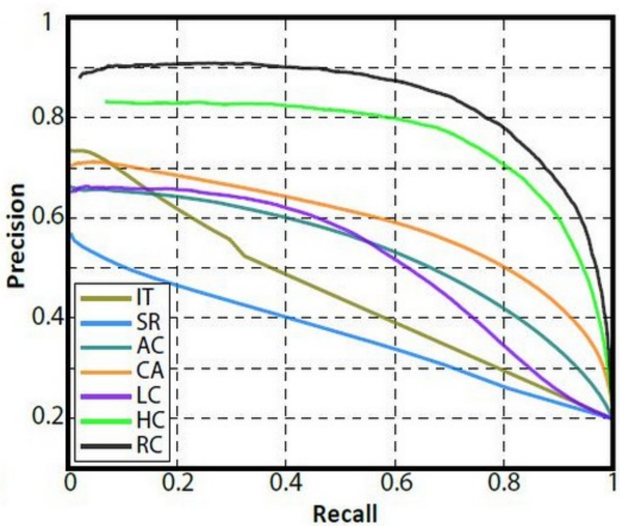
Reliable estimation of visual saliency allows appropriate processing of images without prior knowledge of their content, and thus remains an important step in many computer vision tasks including image segmentation, object recognition, and adaptive compression. We propose a regional contrast based saliency extraction algorithm, which simultaneously evaluates global contrast differences and spatial coherence. The proposed algorithm is simple, efficient, and yields full resolution saliency maps. Our algorithm consistently outperformed existing saliency detection methods, yielding higher precision and better recall rates, when evaluated using one of the largest publicly available data sets. We also demonstrate how the extracted saliency map can be used to create high quality segmentation masks for subsequent image processing.

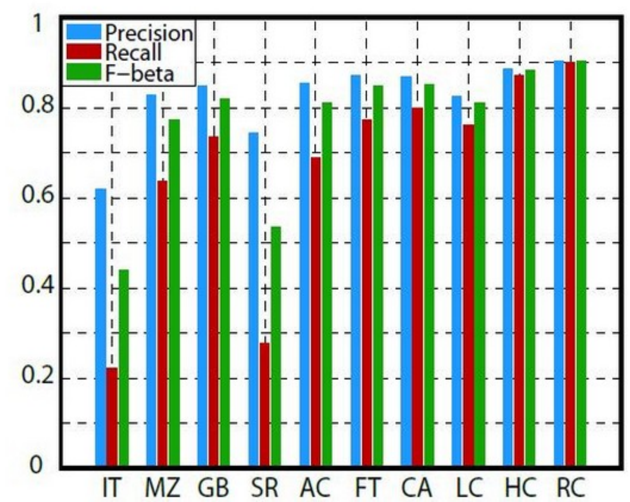
Paper

Ming-Ming Cheng, Guo-Xin Zhang, Niloy J. Mitra, Xiaolei Huang, Shi-Min Hu. Global Contrast based Salient Region Detection. IEEE CVPR, p. 409-416, Colorado Springs, Colorado, USA, June 21-23, 2011. [Project page] [Bib] [Pdf 5MB] [Pdf 15M] [Pdf 中文版] [C++] [Poster] [FAQs]

Comparisons with state of the art methods



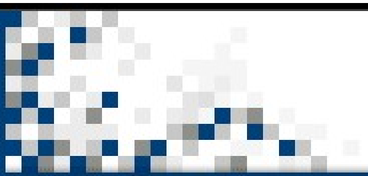




- University home
- University in Finnish
- For Degree Applicants
- For Students
- For Visitors
- For Staff
- Faculties and Departments
- Focus Areas
- Library
- Maps

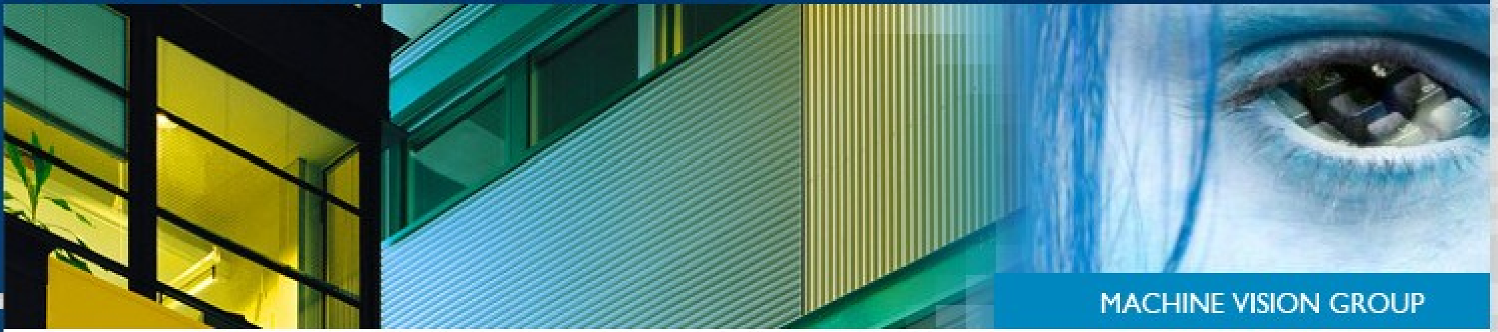
Search

Titles Text



UNIVERSITY of OULU
OULUN YLIOPISTO

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING



Back

Login

- CMV Home
- About CMV
- Research
- LBP Bibliography
- Demos and videos
- CMV in media
- Talks and tutorials
- Publications
- Projects
- Regular courses
- Lectures and seminars
- **Downloads**
- Personnel
- Partners and sponsors
- Links
- Careers
- Contact

Matlab codes for measuring image saliency

Matlab implementation of the saliency measure in [1]

saliencyMeasure.m (2010-06-01, ver 0.1) See function help for instructions and examples.

Publications:

[1] *Rahtu E & Kannala J & Salo M & Heikkilä J (2010)* **Segmenting salient objects from images and Videos**. Proc. European Conference on Computer Vision (ECCV 2010) (Full paper, Stability analysis)

[2] *Rahtu E & Heikkilä J (2009)* **A simple and efficient saliency detector for background subtraction**. Proc. International Workshop on Visual Surveillance (VS2009), 1137-1144. (Full paper)

If you encounter problems or find bugs in the implementation, please contact Esa Rahtu (erahtu at ee.oulu.fi).

Essa Rahtu

CMV/Downloads/saliency (last edited 2011-11-17 14:11:53 by WebMaster)

Immutable Page Info Attachments More Actions:

Yuming Fang's HomePage >

Saliency Detection in the Compressed Domain for Adaptive Image Retargeting

Abstract:

Saliency detection plays important roles in many image processing applications such as Regions of Interest (ROI) extraction and image resizing. Existing saliency detection models are built in the uncompressed domain. Since most images ove
domain such as JPEG (Joint Photographic Experts Group), we propose a novel saliency detection model in the compressed domain in this paper. The intensity, color and texture features of the image are extracted from DCT coefficients in the
obtained based on the Hausdorff distance calculation and feature map fusion. Based on the proposed saliency detection model, we further design an adaptive image retargeting algorithm in the compressed domain. The proposed image retargeti
composed of the block-based seam carving and the image scaling to resize images. A new definition of texture homogeneity is given to determine the amount of removal block-based seams. Thanks to the directly derived accurate saliency info
image retargeting algorithm effectively preserves the visually important regions for images, efficiently removes the less crucial regions, and therefore significantly outperforms the relevant state-of-the-art algorithms, as demonstrated with the in-

Paper:

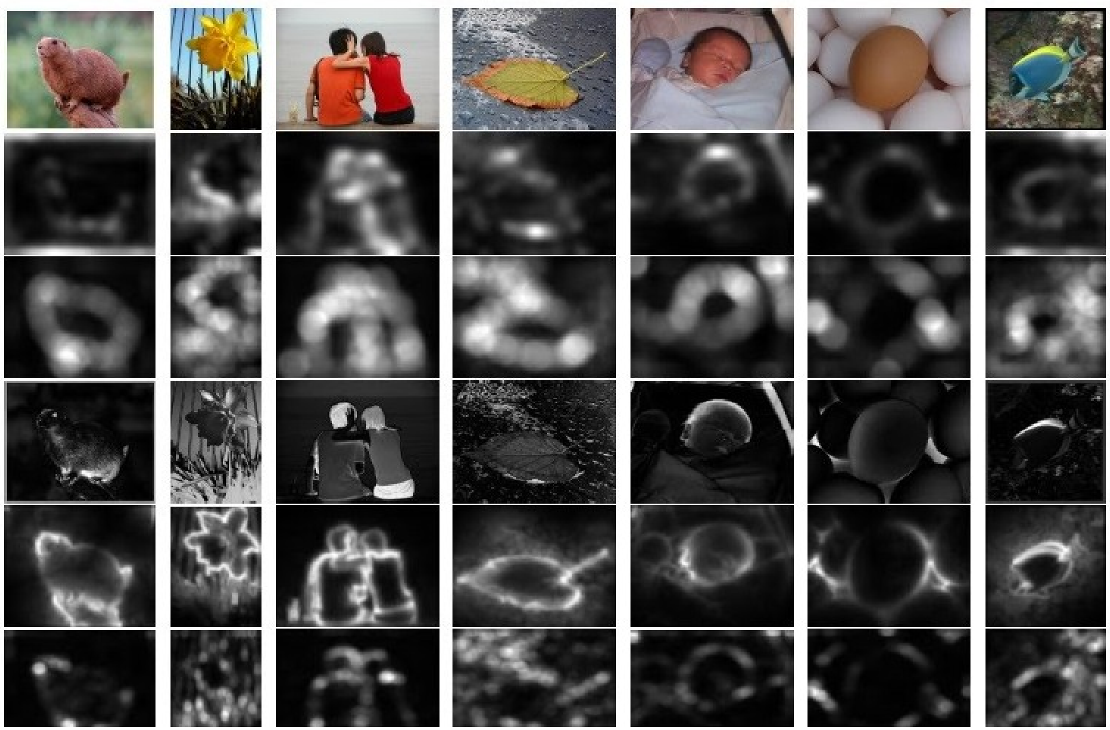
Yuming Fang, Zhenzhong Chen, Weisi Lin, Chia-Wen Lin, 'Saliency Detection in the Compressed Domain for Adaptive Image Retargeting', Accepted by Transactions on Image Processing.(will come out soon!)
Yuming Fang, Zhenzhong Chen, Weisi Lin, Chia-Wen Lin, 'Saliency-based Image Retargeting in the Compressed Domain', *ACM International Conference on Multimedia 2011 (ACM MM11).* [pdf]

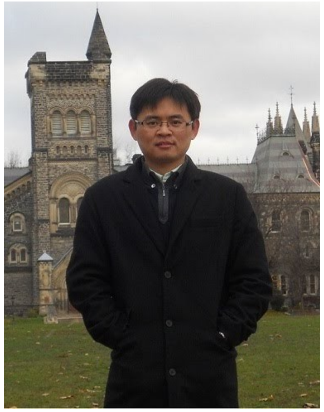
Download:

Matlab Code:
(1) We released the Matlab code for this project. You can download the code here.

Saliency Map Results from EPFL Database:
(2) The saliency map results for the 1000 images in EPFL database can be downloaded here.

Experimental Results:





Yuming Fang

A Co-saliency Model of Image Pairs

Abstract

In this paper, we introduce a method to detect co-saliency from an image pair that may have some objects in common. The co-saliency is modeled as a linear combination of the single-image saliency map (SIS
term is designed to describe the local attention, which is computed by using three saliency detection techniques available in literature. To compute the MISM, a co-multilayer graph is constructed by dividing the
node in the graph is described by two types of visual descriptors, which are extracted from a representation of some aspects of local appearance, e.g., color and texture properties. In order to evaluate the simila
pair SimRank algorithm to compute the similarity score. Experimental evaluation on a number of image pairs demonstrates the good performance of the proposed method on the co-saliency detection task.

Paper

Hongliang Li, King Ngi Ngan, "A Co-saliency Model of Image Pairs," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3365-3375, 2011. [PDF]

Hongliang Li

Results

SDSP: A Novel Saliency Detection Method by Combining Simple Priors

Lin Zhang, Zhongyi Gu, and Hongyu Li

School of Software Engineering, Tongji University, Shanghai

Introduction

This is the website for our paper "SDSP: A Novel Saliency Detection Method by Combining Simple Priors", *in* Proc. ICIP, *2013*.

Salient regions detection from images is an important and fundamental research problem in neuroscience and psychology and it serves as an indispensible step for numerous machine vision tasks. detection method, namely SDSP, by combining three simple priors. At first, the behavior that the human visual system detects salient objects in a visual scene can be well modeled by band-pass fi center of an image. Thirdly, warm colors are more attractive to people than cold colors are. Extensive experiments conducted on the benchmark dataset indicate that SDSP could outperform the o accuracy. Moreover, SDSP has a quite low computational complexity, rendering it an outstanding candidate for time critical applications. The following figure shows the flowchart for SDSP comp

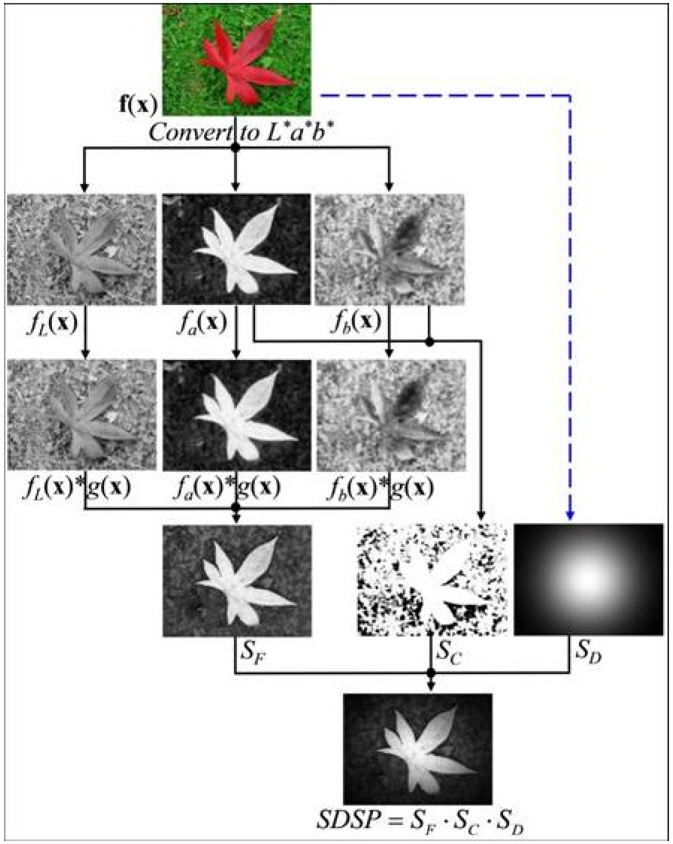


Illustration for the computation process of SDSP.

Lin Zhang

Visual Saliency Based on Scale-Space Analysis in the Frequency Domain

Jian Li

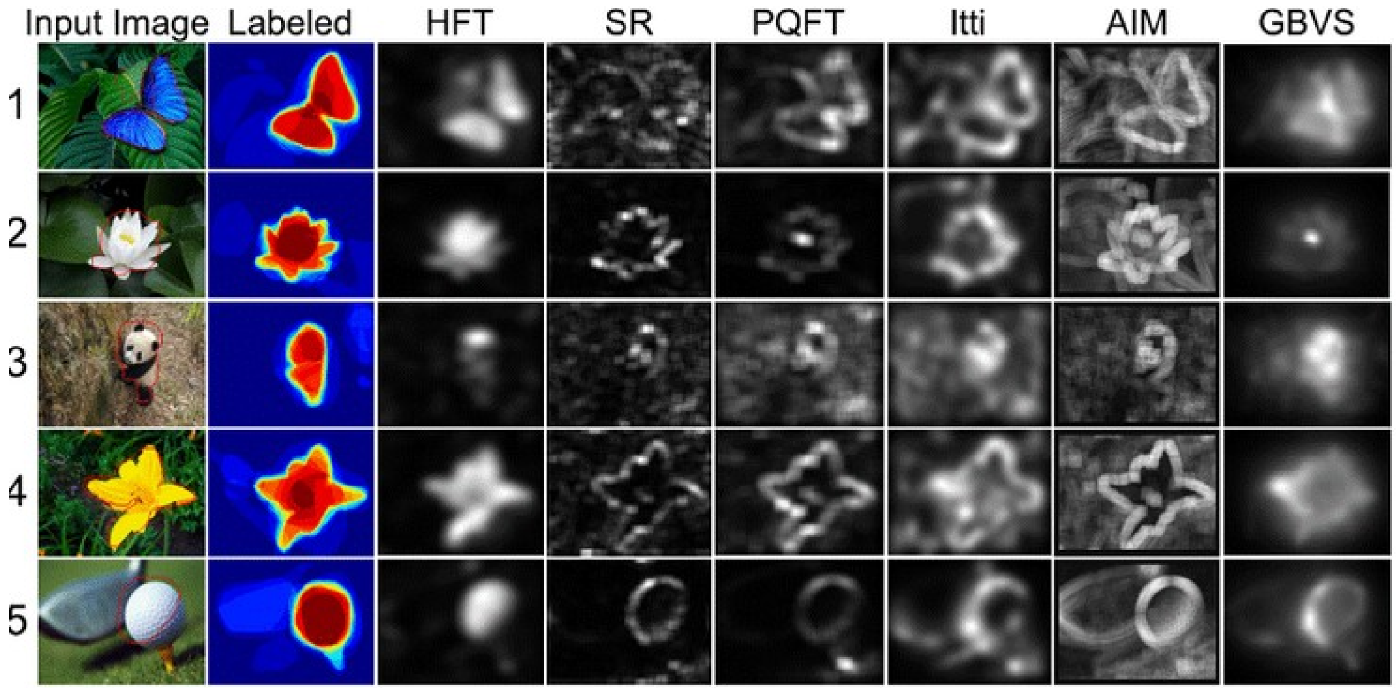
IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI) In press

[PDF] [Matlab code] is available now. The evaluation code can be found here.

If you have any question about this paper, please feel free to contact the first author (Jian Li). If you use this code, please cite our paper.

Some Experimental Results (**see the complete results in the bottom**)

The proposed model (HFT) is compared with SR,PQFT, Itti's model, AIM, GBVS and Human labeled salient results.



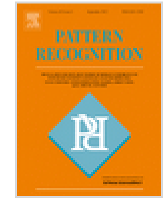
(A) Response to images with large salient regions

ELSEVIER

Pattern Recognition

Volume 45, Issue 9, September 2012, Pages 3114–3124

Best Papers of Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'2011)



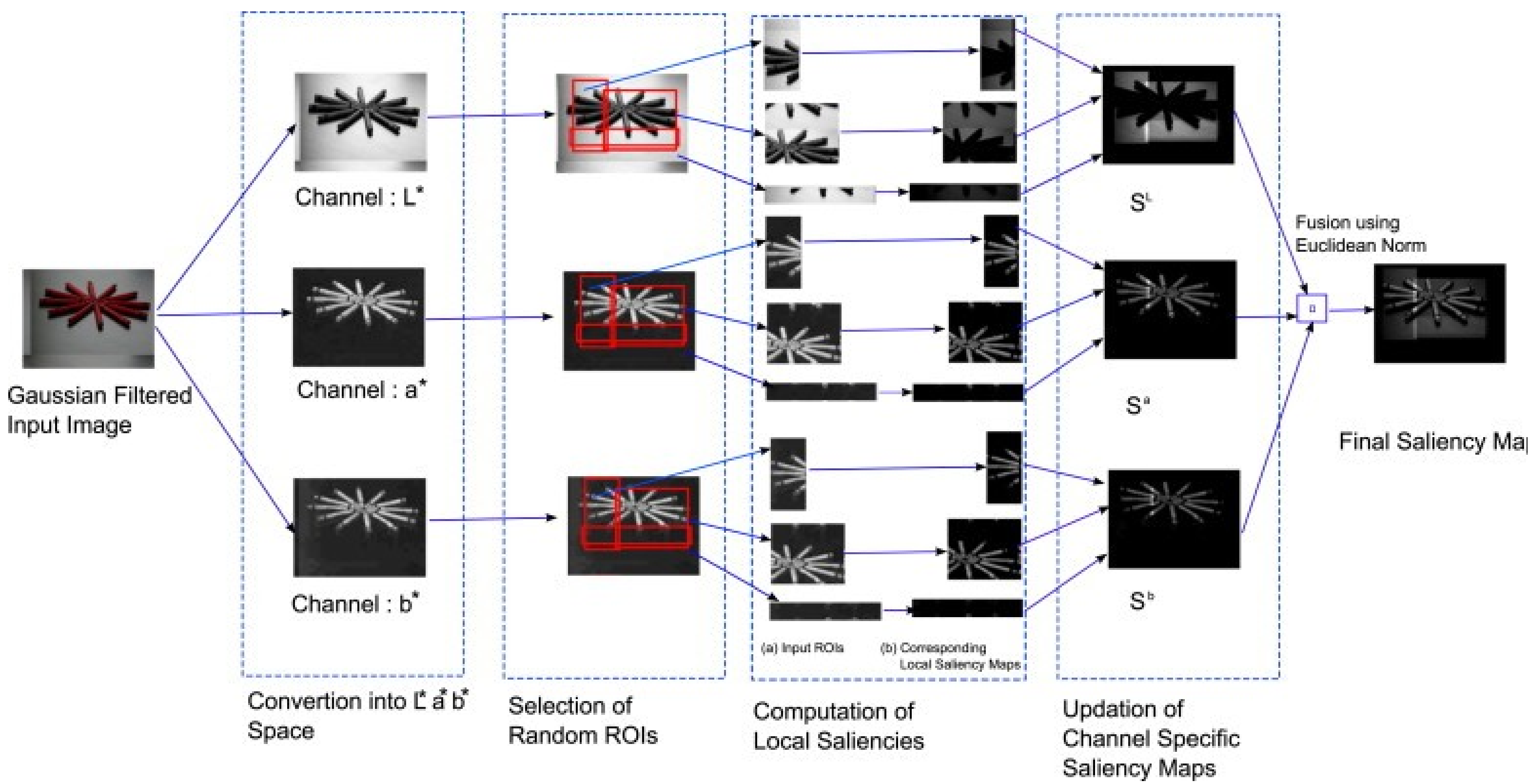
Vikram

A saliency map based on sampling an image into random rectangular regions of interest

Tadmeri Narayan Vikram[a, b], Marko Tscherepanow[a], Britta Wrede[a, b]

[a] Applied Informatics Group, Bielefeld University, Bielefeld 33615, Germany

[b] Research Institute for Cognition and Robotics (CoR Lab), Bielefeld University, Bielefeld 33615, Germany



Salience Distance Transform

Paul Rosin

The distance transform efficiently computes the distance at each pixel to the nearest feature pixel. It is a useful tool in computer vision, and has many applications in model matching, as well as extracting features

A weakness of the approach is the requirement to provide a binary feature map. Thus, when applying it to edge maps for instance, this means that the edges must be thresholded. Over or under thresholding then

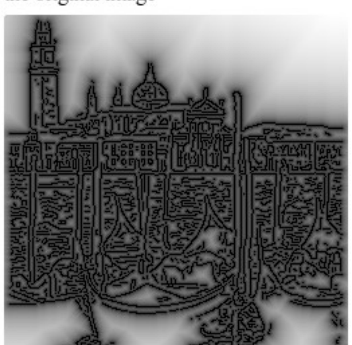
We have developed an improvement called the **salience distance transform** which avoids thresholding, and instead combines other properties of the features.

The following examples show how different thresholds produce very different results from the standard distance transform. Even without thresholding the salience distance transform reduces the effect of the low level edge clutter by incorporating edge salient property further accentuates the dominant features in the scene.

the original image

edge map

log mapped (standard) distance transform of thresholded edges



log mapped (standard) distance transform of edges when thresholded at a different level

salience distance transform using edge magnitude (no thresholding)

salience distance transform using edge magnitude and edge list len

More details are given in:

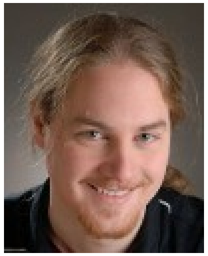
- P.L. Rosin and G.A.W. West, "Salience distance transforms", CVGIP: Graphical Models and Image Processing, vol 57, no. 6, pp. 483-521, 1995.

You can download code to implement the salience distance transform.

Center-surround Divergence of Feature Statistics for Salient Object Detection

Dominik A. Klein and Simone Frintrop

ICCV 2011



Dominik A. Klein

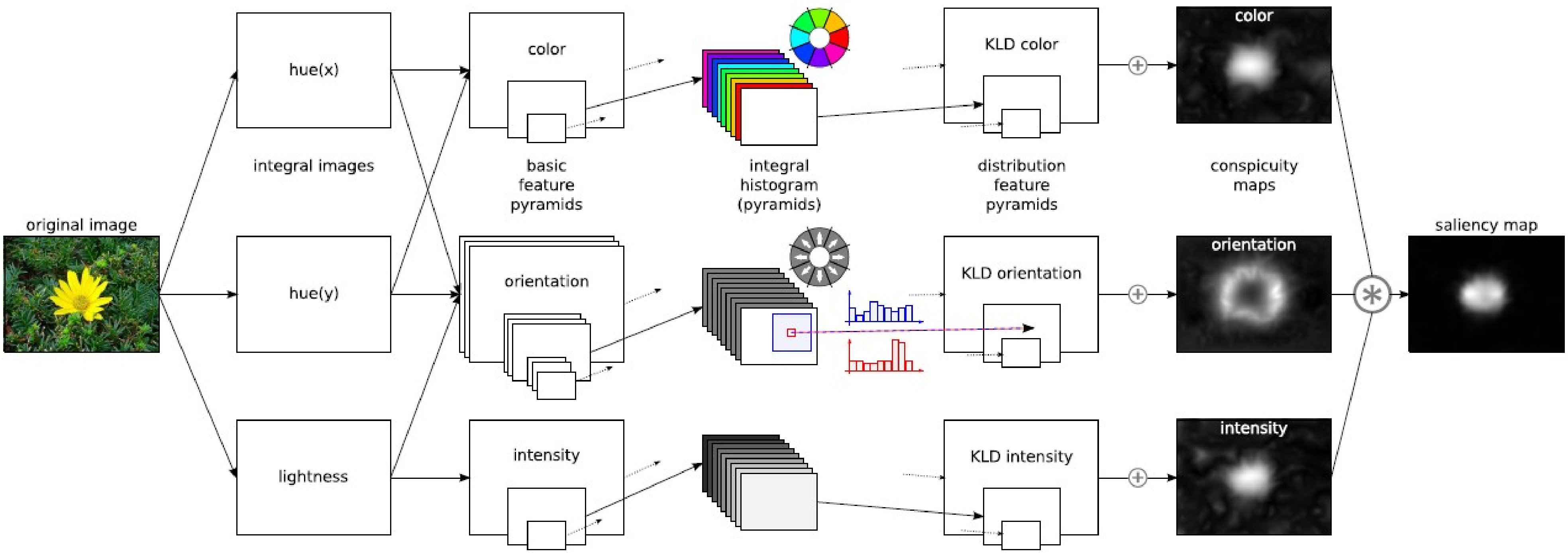
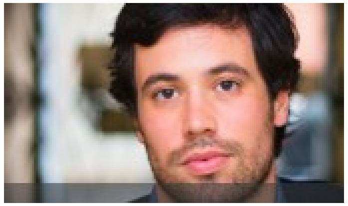


Figure 2. Schematic overview of our saliency system BITS.

Saliency Filters: Contrast Based Filtering for Salient Region Detection



Federico Perazzi

Saliency Filters: Contrast Based Filtering for Salient Region Detection

[1]Federico Perazzi [2]Philipp Krähenbül [1]Yael Pritch [1]Alexander Hornung
[1]Disney Research Zurich [2]Stanford University

(a) Source image. (b) Abstraction. (c) Uniqueness. (d) Distribution. (e) Saliency.

Illustration of the main phases of our algorithm. The input image is first abstracted into perceptually homogeneous elements. Each element is represented by the mean color of the pixels belonging to it. We then define two contrast measures per element based on the uniqueness and spatial distribution of elements. Finally, a saliency value is assigned to each pixel.

Abstract

Saliency estimation has become a valuable tool in image processing. Yet, existing approaches exhibit considerable variation in methodology, and it is often difficult to attribute improvements in result quality to specific algorithm properties. In this paper we reconsider some of the design choices of previous methods and propose a conceptually clear and intuitive algorithm for contrast-based saliency estimation. Our algorithm consists of four basic steps. First, our method decomposes a given image into compact, perceptually homogeneous elements that abstract unnecessary detail. Based on this abstraction we compute two measures of contrast that rate the uniqueness and the spatial distribution of these elements. From the element contrast we then derive a saliency measure that produces a pixel-accurate saliency map which uniformly covers the objects of interest and consistently separates fore- and background. We show that the complete contrast and saliency estimation can be formulated in a unified way using high- dimensional Gaussian filters. This contributes to the conceptual simplicity of our method and lends itself to a highly efficient implementation with linear complexity. In a detailed experimental evaluation we analyze the contribution of each individual feature and show that our method outperforms all state-of-the-art approaches.

Left to right: input images, abstraction into homogeneous elements, our saliency computation, ground truth labeling.

Combined models

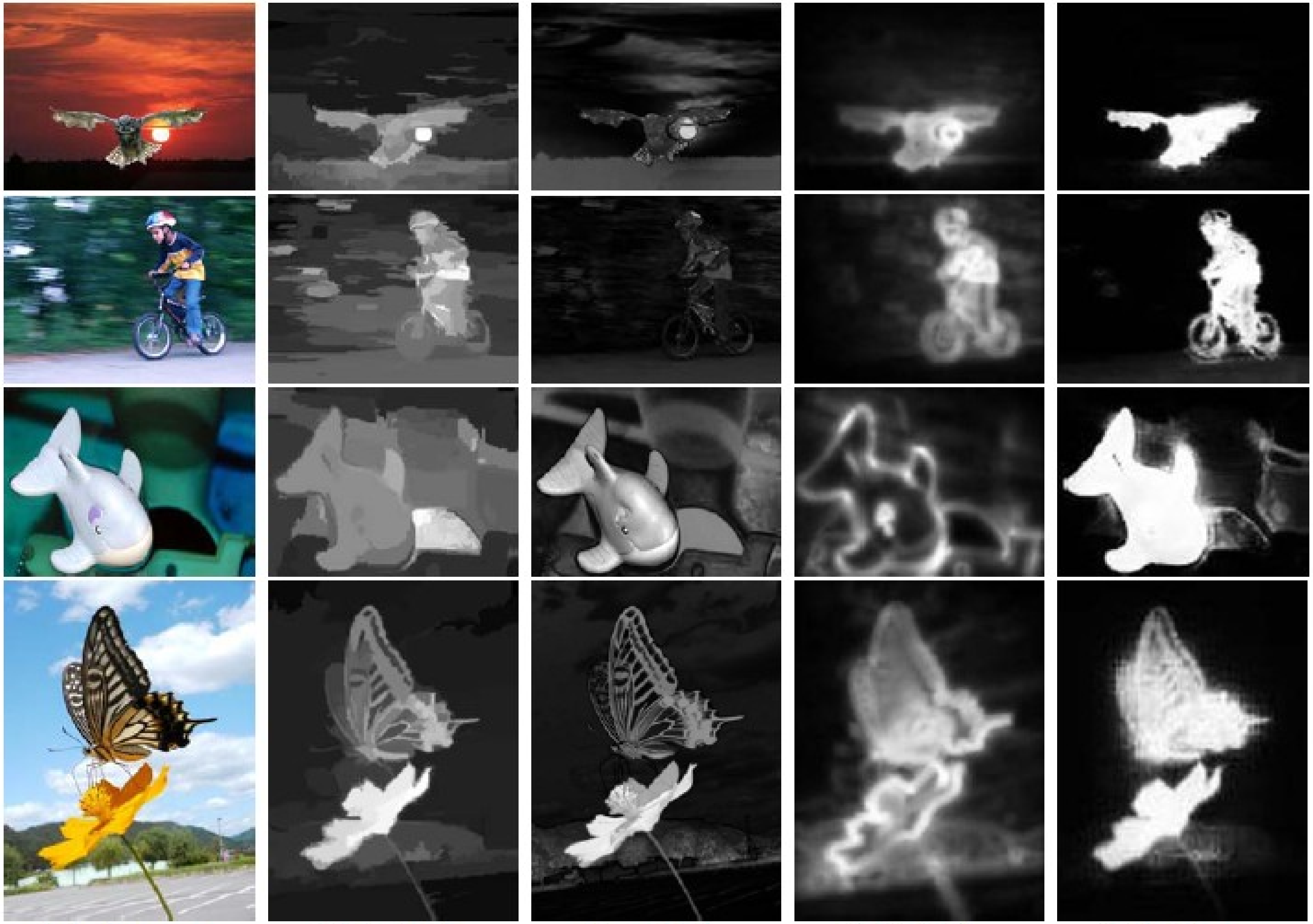
- Building a strong model out of weak ones
- Borji et al., ECCV 2012
- Inspired Long Mai to use CRF for saliency aggregation; CVPR 2013

$$p(x_f|M_1, M_2, \cdots, M_K) \propto \frac{1}{Z} \prod_{k=1}^{K} p(x_f|M_k)$$

$$p(x_f|M_1, M_2, \cdots, M_K) \propto \frac{1}{Z} \sum_{k=1}^{K} \mathcal{G}\big(p(x_f|M_k)\big)$$

Saliency Aggregation: A Data-driven Approach

Long Mai Yuzhen Niu Feng Liu
Department of Computer Science, Portland State University
Portland, OR, 97207 USA CVPR 2013



(a) Image (b) GC (c) FT (d) CA (e) Ours

Plan

- Models
- Challenges
  - Datasets
  - Scores
  - Center-Bias (CB)
- Benchmark
- Summary and Future

Datasets

- Spatial
  - MIT
  - Bruce and Tsotsos (Torronto)
  - NUSEF
  - Kootstra
  - FIFA (Cerf)
- Spatio-Temporal
  - CRCNS
  - DIEM

OVERVIEW OF EYE TRACKING DATASETS

Stefan Winkler and Ramanathan Subramanian

Advanced Digital Sciences Center (ADSC), University of Illinois at Urbana-Champaign, Singapore

Our Databases

1. ASCMN database 24 videos (5 classes: Abnormal, Surveillance, Crowd, Moving, Noise) / 10 observers per video. No specific task was asked to the viewers. The eye-tracker raw data sampled at the frame rate of each video are provided. If you want to compute the score (KL divergence, NSS and AUROC) of your own dynamic saliency model, you can do it on our benchmark page.
   This work was carried out in collaboration with Dubravko Culibrk. If you use it, please cite this reference:

   > *N. RICHE, M. MANCAS, D. ĆULIBRK, V. ĆRNOJEVIC, B. GOSSELIN, T. DUTOIT, 2012, "Dynamic saliency models and human attention: a comparative study on videos", Proceedings of the 11th Asian Conference on Computer Vision (ACCV), Daejeon, Korea, November 5-9.*

2. Mouse tracking database 91 images (3 classes: natural images, advertisements, websites) / 40 - 60 observers per image. The task was to show with the mouse the regions of the image which are attended. In this case mouse and eye tracking have high correlation (see: here)
   This database was used during Matei's PhD thesis in 2007. If you use it, please cite this reference :

   > *M. MANCAS, 2009, "Relative Influence of Bottom-Up and Top-Down Attention", Attention in Cognitive Systems, Lecture Notes in Computer Science, Springer Berlin / Heidelberg, ISSN 0302-9743 (Print) 1611-3349 (Online), Volume 5395/2009, DOI 10.1007/978-3-642-00582-4, ISBN 978-3-642-00581-7, February 2009.*

Online available databases

Here, you can find a recent state of the art of online available databases.

Still Images

Fixation ground truth (by eye tracker)

1. AIM : 120 images viewed by 20 users (Neil D. B. Bruce and John K. Tsotsos 2005).
2. LeMeur : 27 images viewed by 40 users (O. Le Meur, P. Le Callet, D. Barba and D. Thoreau 2006).
3. Torralba : 73 images viewed by 24 users (Antonio Torralba, Aude Oliva, Monica Castelhano and John Henderson 2006).
4. Kootstra : 100 images viewed by 31 users (Kootstra, G., Nederveen, A. and de Boer, B. 2008).
5. DOVES : 101 images viewed by 29 users (van der Linde, I., Rajashekar, U., Bovik, A.C., Cormack, L.K. 2009).
6. Ehinger : 912 images viewed by 14 users (Krista A. Ehinger, Barbara Hidalgo-Sotelo, Antonio Torralba and Aude Oliva 2009).
7. FIFA : 180 images viewed by 8 users (Moran Cerf, Jonathan Harel, Wolfgang Einhäuser and Christof Koch 2009).
8. Judd : 1003 images viewed by 15 users (Tilke Judd, Krista Ehinger, Frédo Durand and Antonio Torralba 2009).
9. NUSEF : 758 images viewed by 75 users (R. Subramanian, H. Katti, N. Sebe1, M. Kankanhalli and T-S. Chua 2010).

Region ground truth (human labeled)

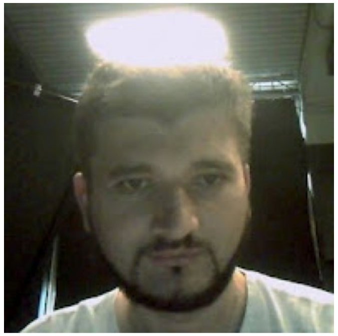
1. MSRA Salient Object Database : 20000 images labeled by 3 users (Tie Liu, Jian Sun, Nan-Ning Zheng, Xiaoou Tang and Heung-Yeung Shum 2007).
2. MSRA Salient Object Database : 5000 images labeled by 9 users (Tie Liu, Jian Sun, Nan-Ning Zheng, Xiaoou Tang and Heung-Yeung Shum 2007).
3. Achanta : 1000 images labeled by 9 users (R. Achanta, S. Hemami, F. Estrada and S. Süsstrunk 2009).

Fixation and region ground truth

1. IRCCyN IVC Berkeley Eyetracker : 80 images labeled and viewed by 25 users (Junle Wang, Damon M. Chandler and Patrick Le Callet 2010).
2. JianLi : 235 images labeled and viewed by 19 users (Jian Li, Martin D. Levine, Xiangjing An and Hangen He 2011).

ImageSal dataset

Videos



Matei Mancas

Video datastes

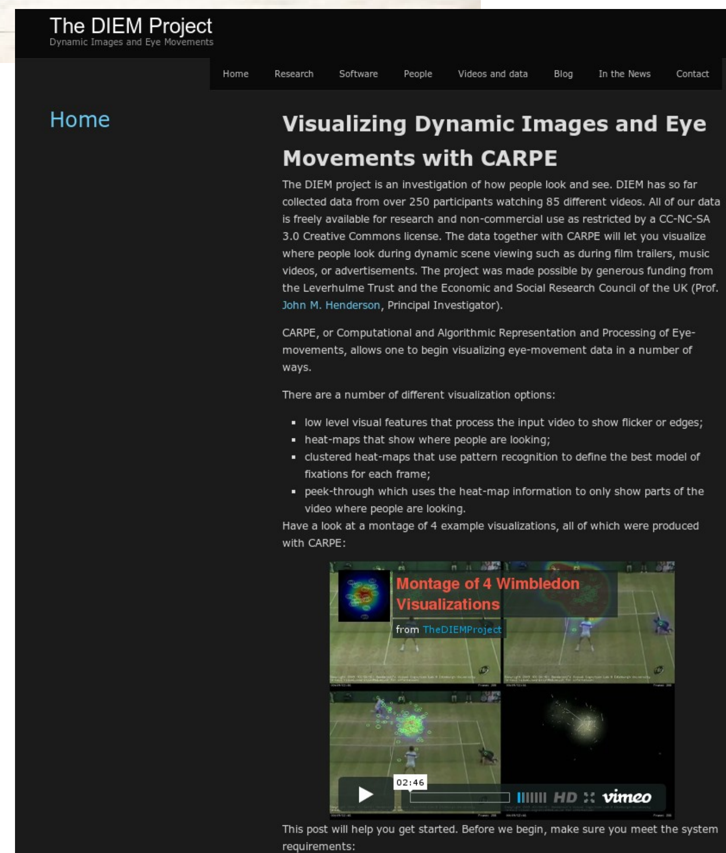
Videos

Fixation ground truth (by eye tracker)

1. CRCNS - Collaborative Research in Computational Neuroscience : 50 videos viewed by 8 users (L. Itti 2006).
2. IRCCyN IVC Eyetracker SD 2008 11 : 51 videos viewed by 37 users (Fadi Boulos, Wei Chen, Benoit Parrein and Patrick Le Callet 2008).
3. The DIEM Project : 85 videos viewed by 250 users (John M. Henderson, Robin Hill,Tim Smith and K. Mital 2009).
4. Lubeck University Dataset : 54 videos viewed by 18 users (Michael Dorr, Thomas Martinetz, Karl Gegenfurtner and Erhardt Barth 2010).
5. Actions in the Eye : 497 107 frames viewed by 16 users (Stefan Mathe and Cristian Sminchisescu 2012).
6. Eye tracking database for standard video sequences : 12 Videos viewed by 15 users (H. Hadizadeh, M. J. Enriquez, and I. V. Bajić 2012).

Marat et al., 2009
Jia Li et al., 2010
Peters and Itti, 2007
Borji et al., 2012
Shic and Scassellati, 2007
Le Meur et al., 2007

John Henderson



The DIEM Project
Dynamic Images and Eye Movements

Home | Research | Software | People | Videos and data | Blog | In the News | Contact

Home

Visualizing Dynamic Images and Eye Movements with CARPE

The DIEM project is an investigation of how people look and see. DIEM has so far collected data from over 250 participants watching 85 different videos. All of our data is freely available for research and non-commercial use as restricted by a CC-NC-SA 3.0 Creative Commons license. The data together with CARPE will let you visualize where people look during dynamic scene viewing such as during film trailers, music videos, or advertisements. The project was made possible by generous funding from the Leverhulme Trust and the Economic and Social Research Council of the UK (Prof. John M. Henderson, Principal Investigator).

CARPE, or Computational and Algorithmic Representation and Processing of Eye-movements, allows one to begin visualizing eye-movement data in a number of ways.

There are a number of different visualization options:

- low level visual features that process the input video to show flicker or edges;
- heat-maps that show where people are looking;
- clustered heat-maps that use pattern recognition to define the best model of fixations for each frame;
- peek-through which uses the heat-map information to only show parts of the video where people are looking.

Have a look at a montage of 4 example visualizations, all of which were produced with CARPE:

This post will help you get started. Before we begin, make sure you meet the system requirements:

CRCNS

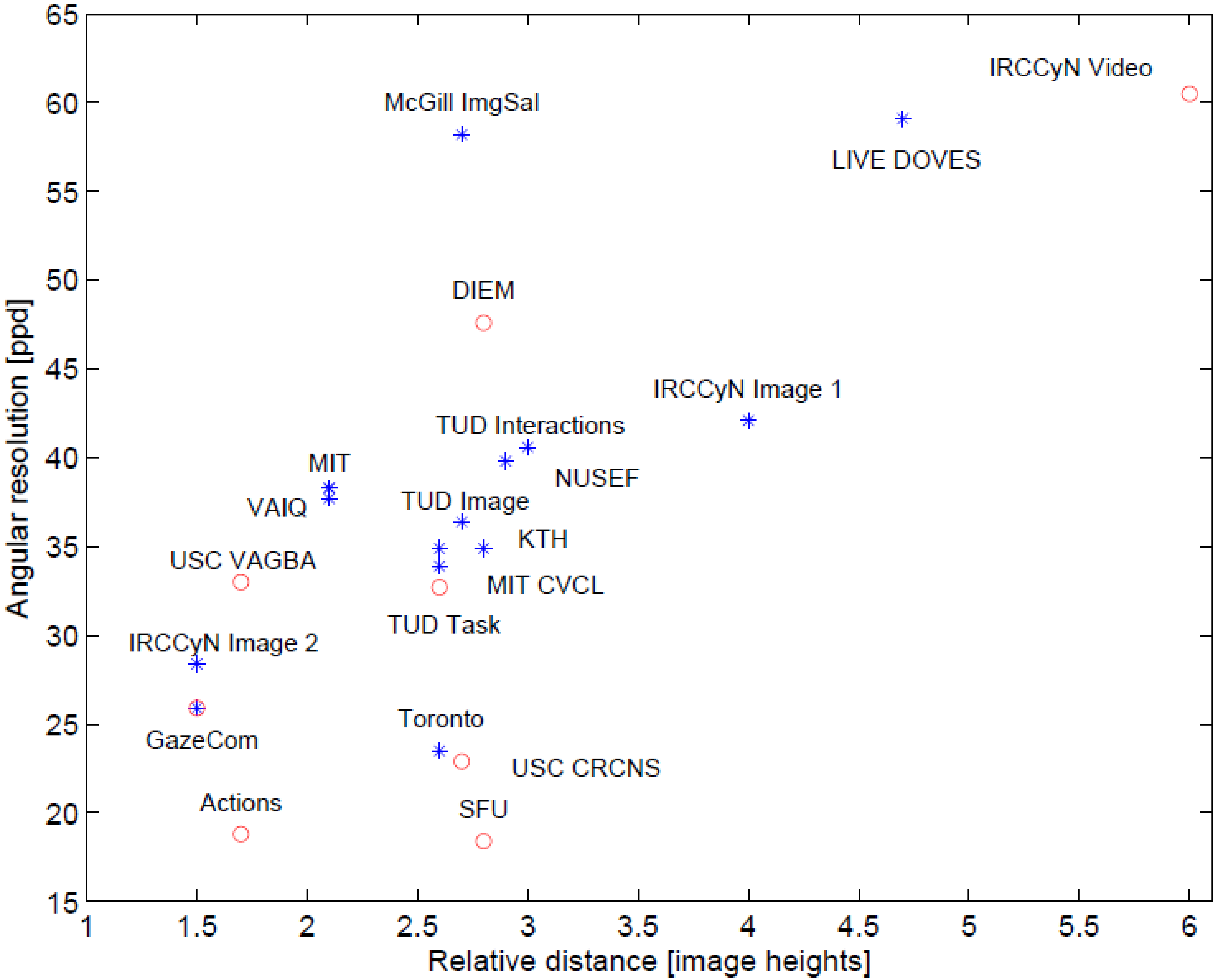
DIEM

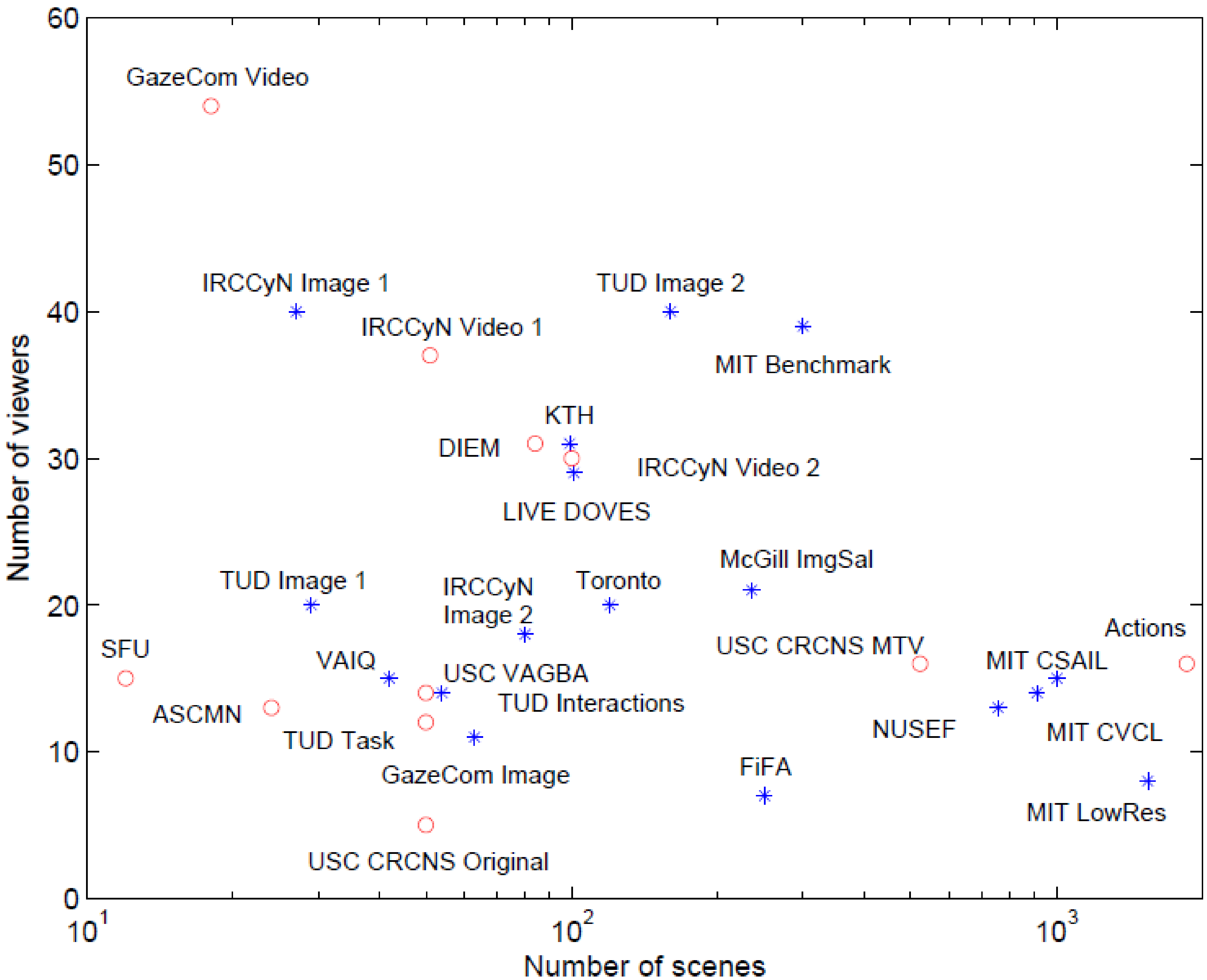

BBC-life-in-cold-blood -1278x710
BBC-wildlife-serpent-1280x704
DIY-SOS-1280x712
advert-bbc4-bees-1024x576
harry-potter-6-trailer-1280x544
music-gummybear-880x720
music-trailer-nine-inch-nails -1280x720
news-tony-blair-resignation -720x540
sport-scramblers-1280x720
sport-wimbledon-federer-final -1280x704
advert-bbc4-library-1024x576
advert-iphone-1272x720
ami-ib4010-closeup-720x576
ami-ib4010-left-720x576
nightlife-in-mozambique-1280x580
one-show-1280x712
pingpong-angle-shot-960x720
pingpong-no-bodies-960x720
tv-uni-challenge-final-1280x712
university-forum-construction-ionic -1280x720

Table 1. Eye tracking datasets at a glance (T is viewing time, D is viewing distance, d is screen diagonal, f is frequency).

Dataset	Year	Type	Scenes	Resolution	Users	Age	T [sec]	D [cm]	d [in]	Screen	Eye Tracker	f [Hz]	Restraint
FiFA [8]	2007	Image	250	1024×768	7		2	80		CRT	EyeLink 1000	1000	Chin rest
GazeCom Image [10]	2010	Image	63	1280×720	11	18-34	2	45	22	CRT	EyeLink II	250	Chin rest
IRCCyN Image 1 [25]	2006	Image	27	≈768×512	40		15			CRT	Cambridge Research	50	
IRCCyN Image 2 [40]	2010	Image	80	481×321	18	19-45	15	40	17	LCD	Cambridge Research	50	
KTH [23]	2011	Image	99	1024×768	31	17-32	5	70	18	CRT	Eyelink I		Headmount
LIVE DOVES [39]	2009	Image	101	1024×768	29	μ =27	5	134	21	CRT	Fourward Tech. Gen. V	200	Bite bar
McGill ImgSal [27]	2013	Image	235	640×480	21			70	17	LCD	Tobii T60	60	
MIT Benchmark [21]	2012	Image	300	≈1024×768	39	18-50	3	61	19		ETL 400 ISCAN	240	Chin rest
MIT CSAIL [22]	2009	Image	1003	≈1024×768	15	18-35	3	61	19				Chin rest
MIT CVCL [11]	2009	Image	912	800×600	14	18-40		75	21	CRT	ISCAN RK-464	240	Head rest
MIT LowRes [20]	2011	Image	1544	1024×860	8	18-55	3	61	19		ETL 400 ISCAN	240	Chin rest
NUSEF [36]	2010	Image	758	1024×860	13	18-35	5	76	17	LCD	ASL	30	
Toronto [5]	2006	Image	120	681×511	20		4	75	21	CRT			
TUD Image 1 [29]	2009	Image	29	varying	20	students	10	70	19	CRT	iView X RED	50	Chin rest
TUD Image 2 [1]	2011	Image	160	600×600	40		8	60	17	CRT	iView X RED	50	Head rest
TUD Interactions [37]	2011	Image	54	768×512	14	22-35		70	17	CRT	SMI	50/60	Chin rest
VAIQ [12]	2009	Image	42	varying	15	20-60	12	60	19	LCD	EyeTech TM3		
Actions [33]	2012	Video	1857	SD	16	21-41	$<$60	60	22	LCD	SMI iView X HiSpeed	500	Chin rest
ASCMN [38]	2012	Video	24	VGA-SD	13	23-35	2-76				faceLAB		
DIEM [34]	2011	Video	85	SD-HD	42	18-36	27-217	90	21		Eyelink 2000	1000	Chin rest
GazeCom Video [10]	2010	Video	18	720p	54	18-34	20	45	22	CRT	EyeLink II	250	Chin rest
IRCCyN Video 1 [35]	2009	Video	51	720×576	37		8-10	276	37	LCD	Cambridge Research	50	
IRCCyN Video 2 [13]	2010	Video	100	720×576	30		10	150	40	LCD	Cambridge Research		
SFU [17]	2012	Video	12	CIF	15	18-30	3-10	80	19	LCD	Locarna Pt-Mini	30	Headmount
TUD Task [2]	2012	Video	50	720p	12	students	20	60	17	CRT	EyeLink II	250	
USC CRCNS Orig. [18]	2004	Video	50	640×480	8	23-32	6-90	80	22	CRT	ISCAN RK-464	240	Chin rest
USC CRCNS MTV [6]	2006	Video	523	640×480	16	23-32	1-3	80	22	CRT	ISCAN RK-464	240	Chin rest
USC VAGBA [28]	2011	Video	50	1080	14	22-32	10	98	46	LCD	ISCAN RK-464	240	Chin rest

Adapted from: OVERVIEW OF EYE TRACKING DATASETS; Winkler and Subramanian

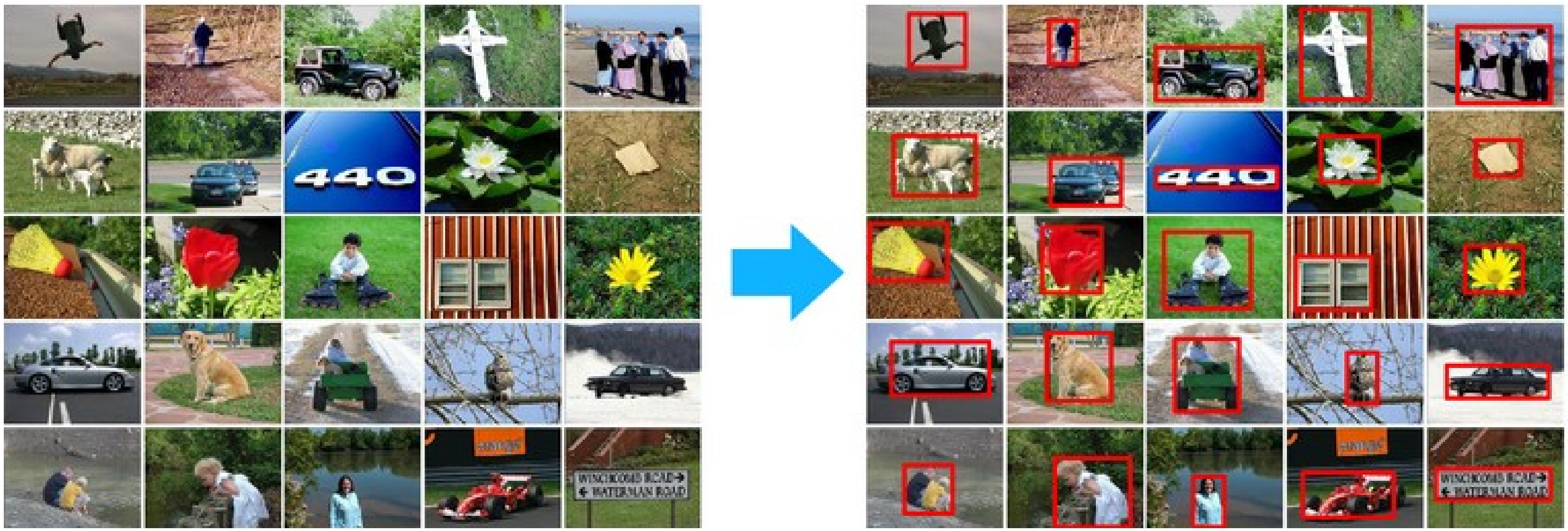

IRCCyN Video
McGill ImgSal
LIVE DOVES
DIEM
IRCCyN Image 1
TUD Interactions
MIT
NUSEF
VAIQ
TUD Image
KTH
USC VAGBA
MIT CVCL
TUD Task
IRCCyN Image 2
GazeCom
Toronto
USC CRCNS
Actions
SFU
Angular resolution [ppd]
Relative distance [image heights]
65
60
55
50
45
40
35
30
25
20
15
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6


GazeCom Video
IRCCyN Image 1
TUD Image 2
IRCCyN Video 1
MIT Benchmark
KTH
DIEM
IRCCyN Video 2
LIVE DOVES
McGill ImgSal
TUD Image 1
IRCCyN
Image 2
Toronto
Actions
SFU
USC CRCNS MTV
VAIQ
MIT CSAIL
USC VAGBA
TUD Interactions
ASCMN
NUSEF
TUD Task
MIT CVCL
FiFA
GazeCom Image
MIT LowRes
USC CRCNS Original
Number of viewers
Number of scenes
0
10
20
30
40
50
60
10^1
10^2
10^3

Salient object detection DBs

- See Borji et al., ECCV 2012
- ASD Achanta et al. (stimuli from MSRA)
- SOD (from Berkley dataset); Movahedi et al.
- SED 1 & SED 2 (Weizmann dataset)
- 5K images from MSRA-B; Huaizu Jiang
- ImageSal dataset
- iCoseg: Interactive cosegmentation by touch

MSRA Salient Object Database



Download:

Image set A --- 20,000 image labeled by three users. (images, labeled rectangles, and readme files)

Image set B --- 5,000 images labeled by nine users. (images, labeled rectangles, and readme files)

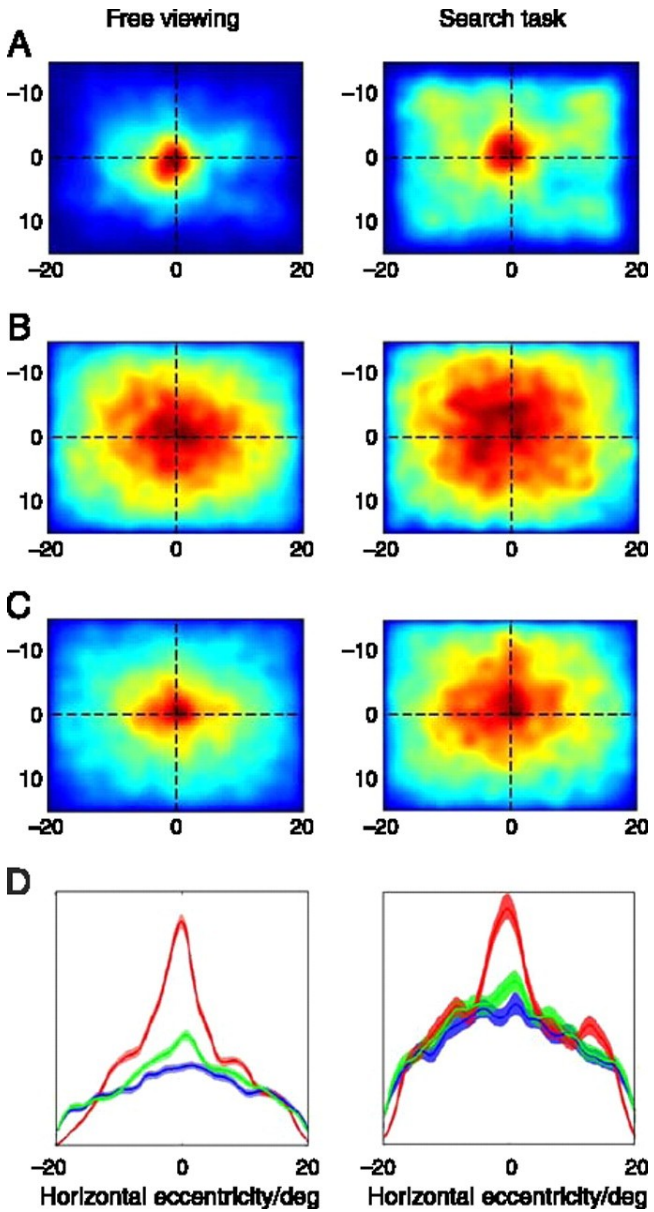
Publication:

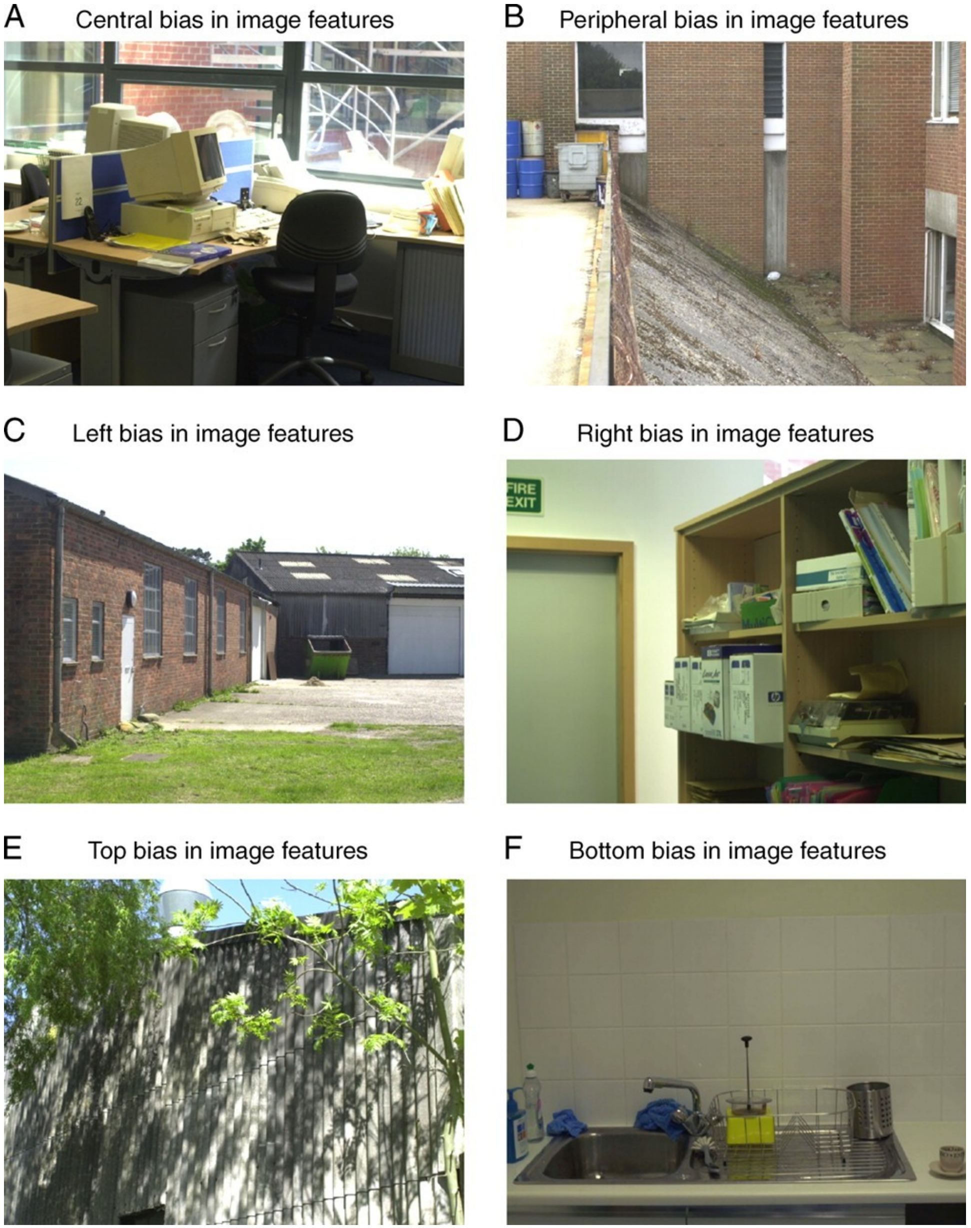
Tie Liu, Jian Sun, Nan-Ning Zheng, Xiaoou Tang and Heung-Yeung Shum. Learning to Detect A Salient Object. In Proc. IEEE Cont. on Computer Vision and pattern Recognition (CVPR), Minneapolis, Minnes

Contactor: jiansun@microsoft.com


ASD
MSRA
SED1
SED2
SOD
Images with smallest objects
Images with largest objects
WARNING

Center-bias

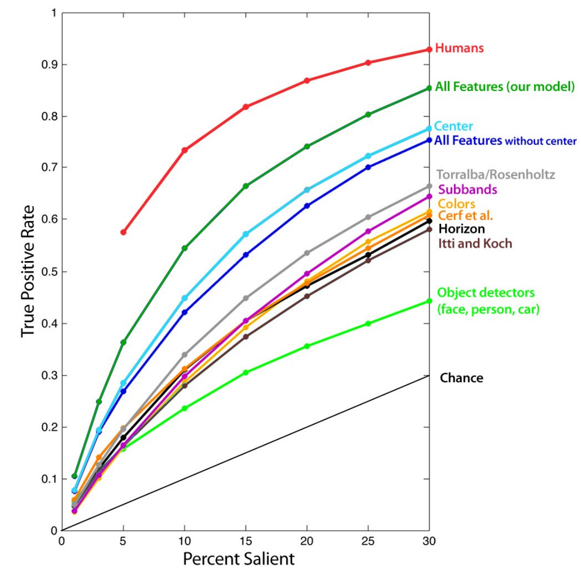


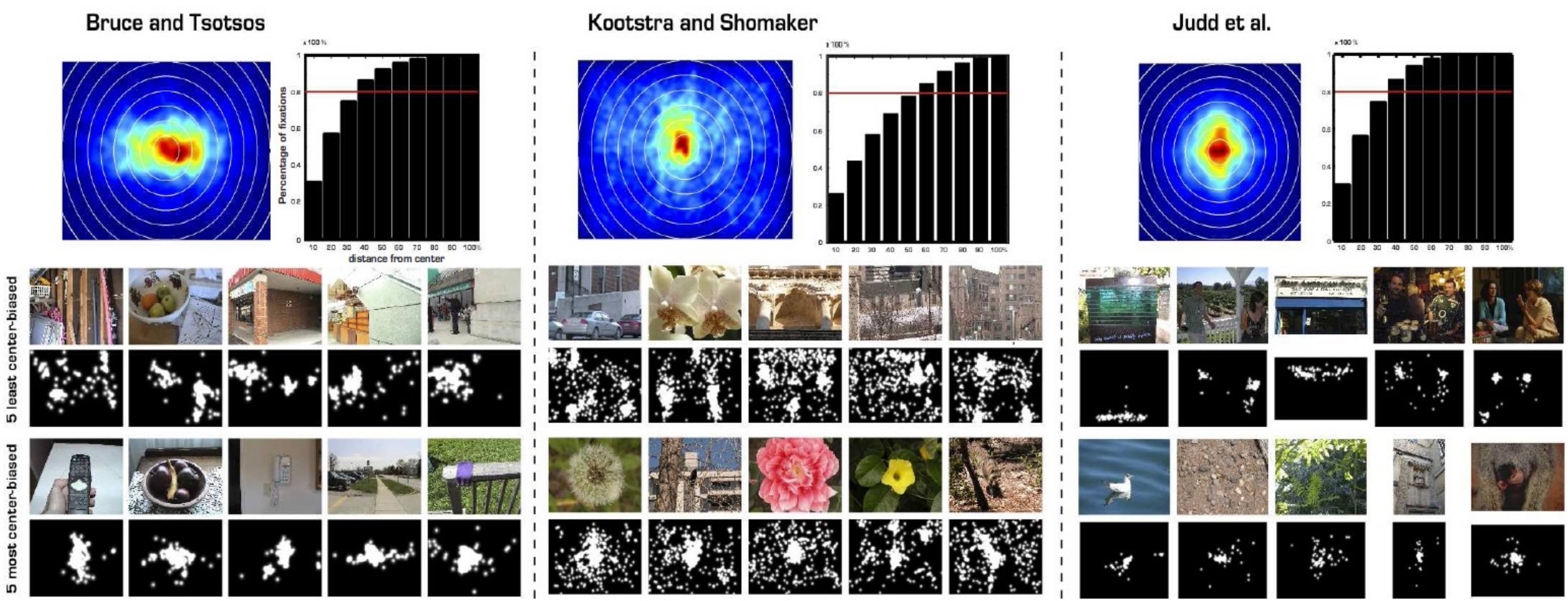


Tatler JOV 2007

Center-bias

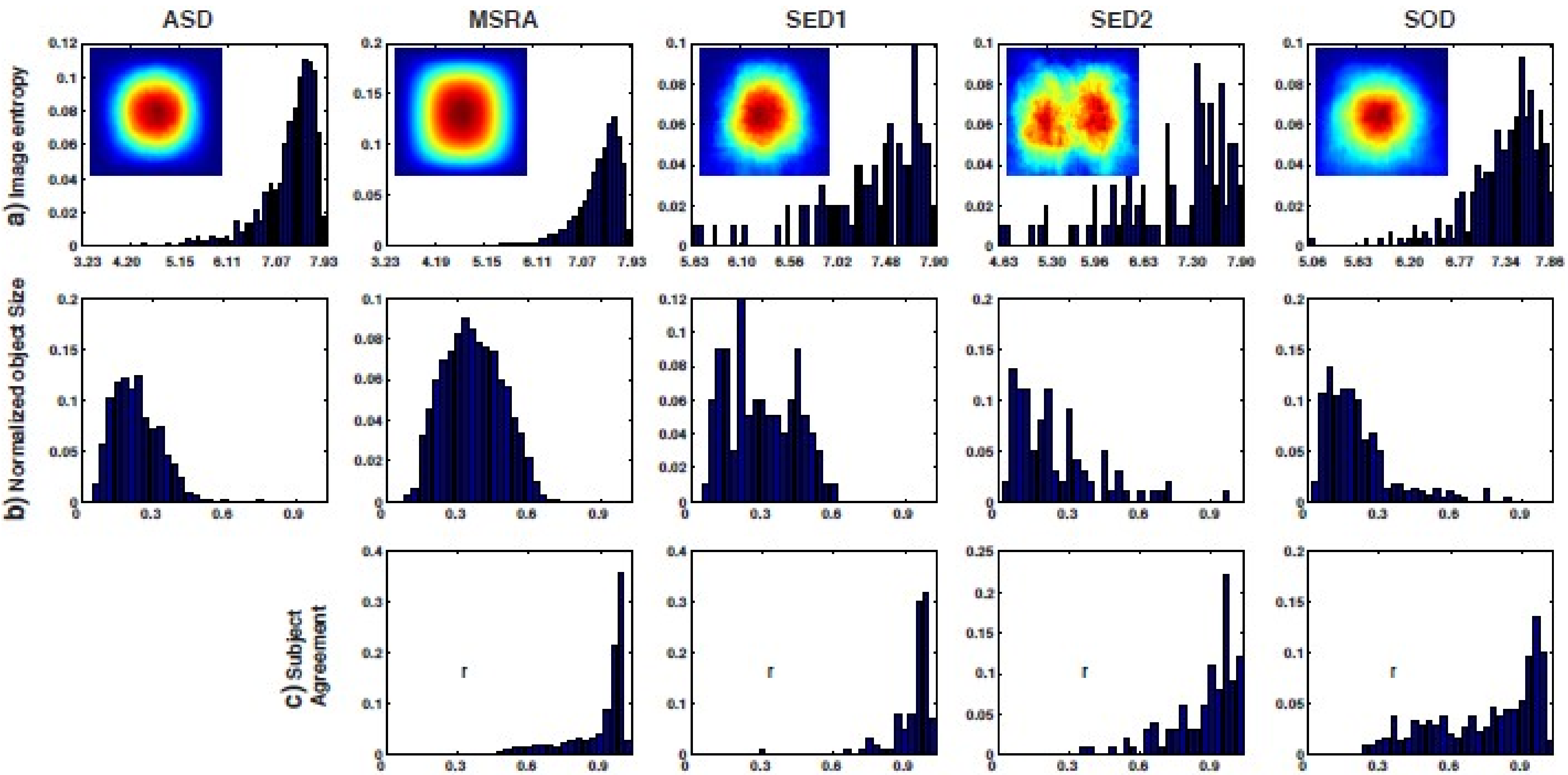
- Viewing strategy
- Photographer bias
- See Tatler et al., 2007, Tseng et al., 2009, Borji et al., VSS 2011





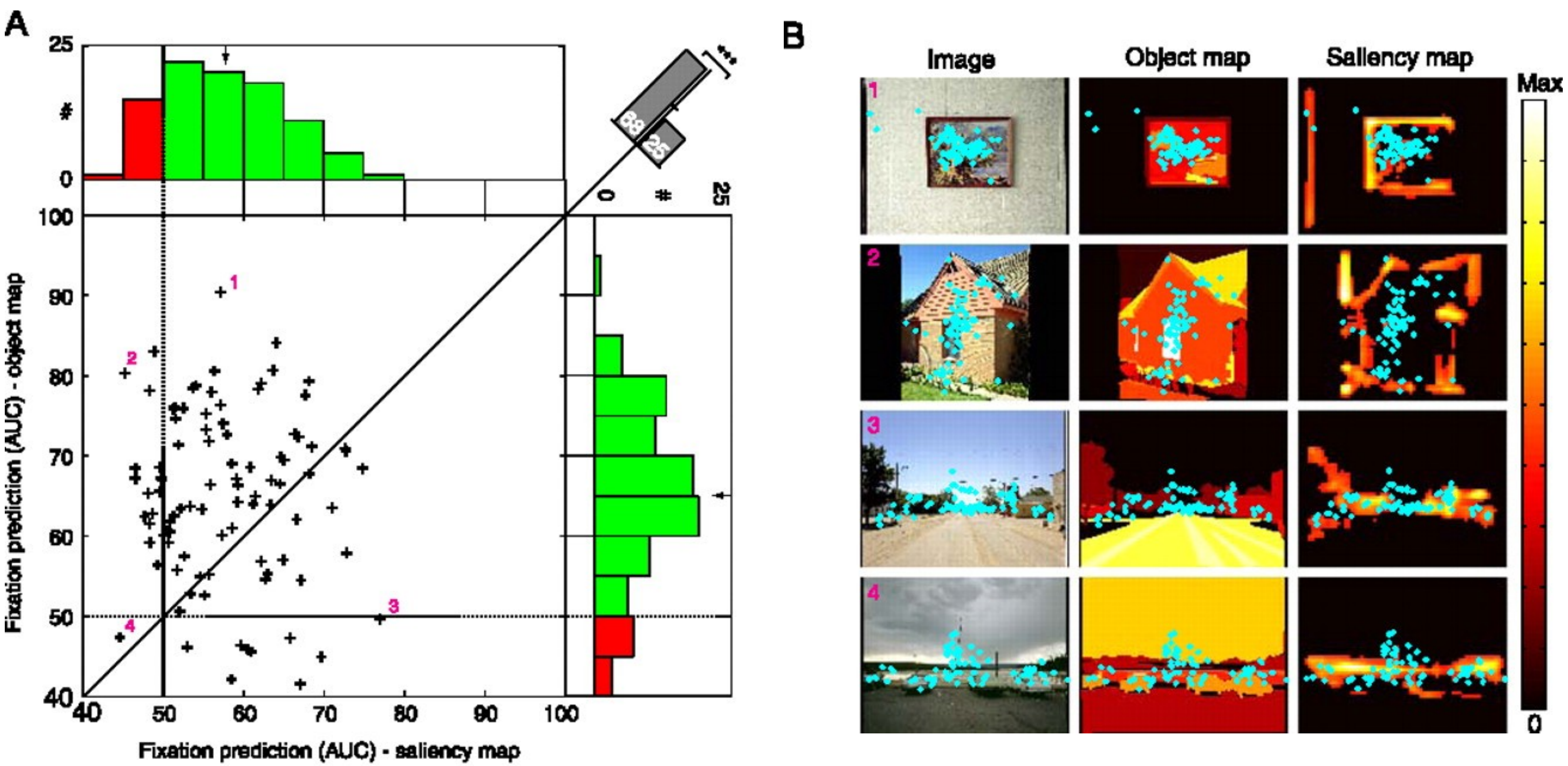
Center-bias

- Salient Object Datasets

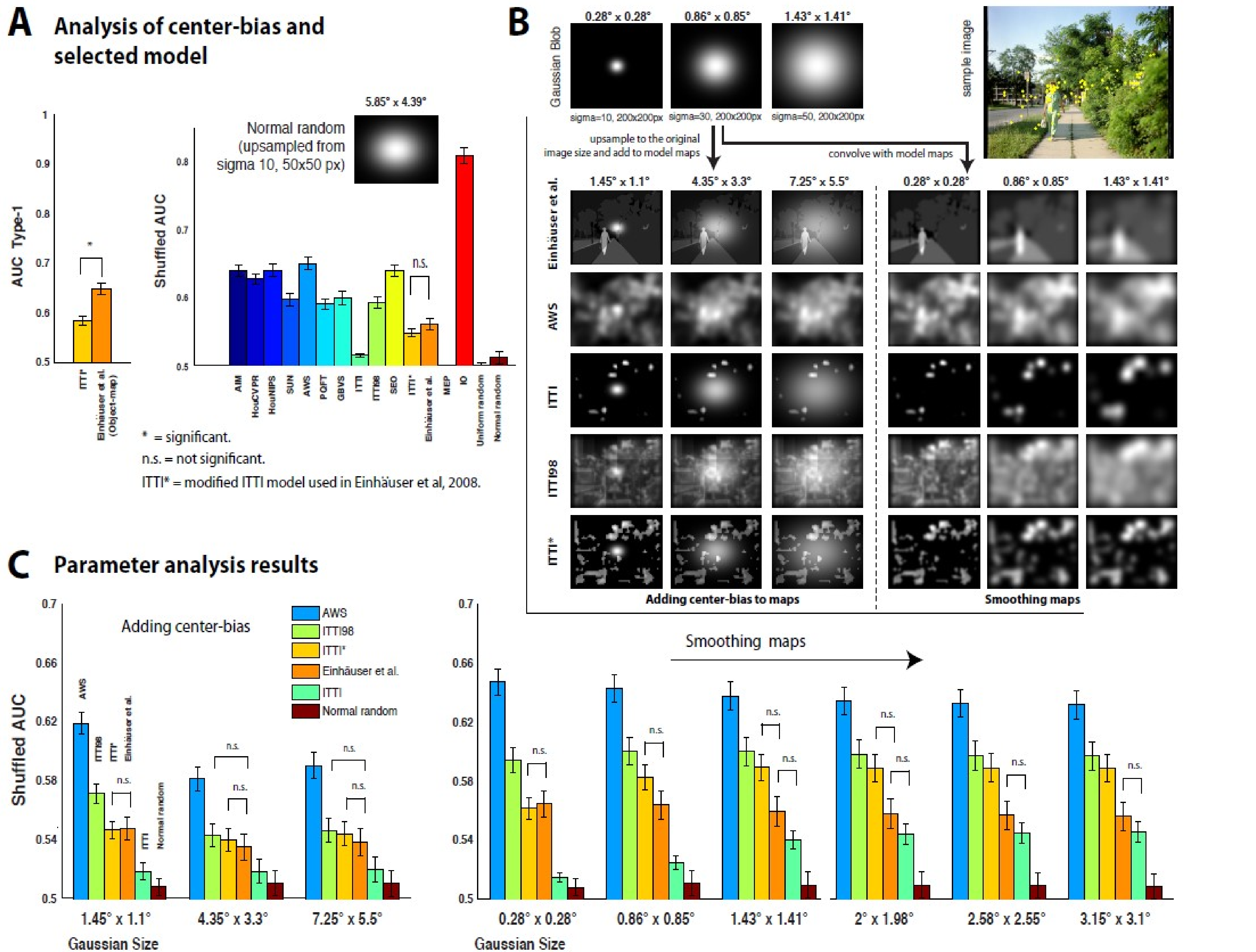


Do objects predict fixations better than early saliency?

- Einhauser et al., JOV 2008



Borji et al., JOV, to appear



Scores

- **Fixation Prediction**
  - AUC (I spotted 4 types of AUC)
  - NSS
  - CC
  - KL
  - EMD
  - Percentile
  - Fixation Saliency Method (FS)
  - String Editing Distances
  - Average Accuracy Score (AAS)
  - Similarity score
- **Saliency Detection**
  - AUC
  - Precision-Recall

Refs:

- Borji and Itti, PAMI 2013
- Peters and Itti, TAP 2008
- Alsam and Sharma, SCIA 2013

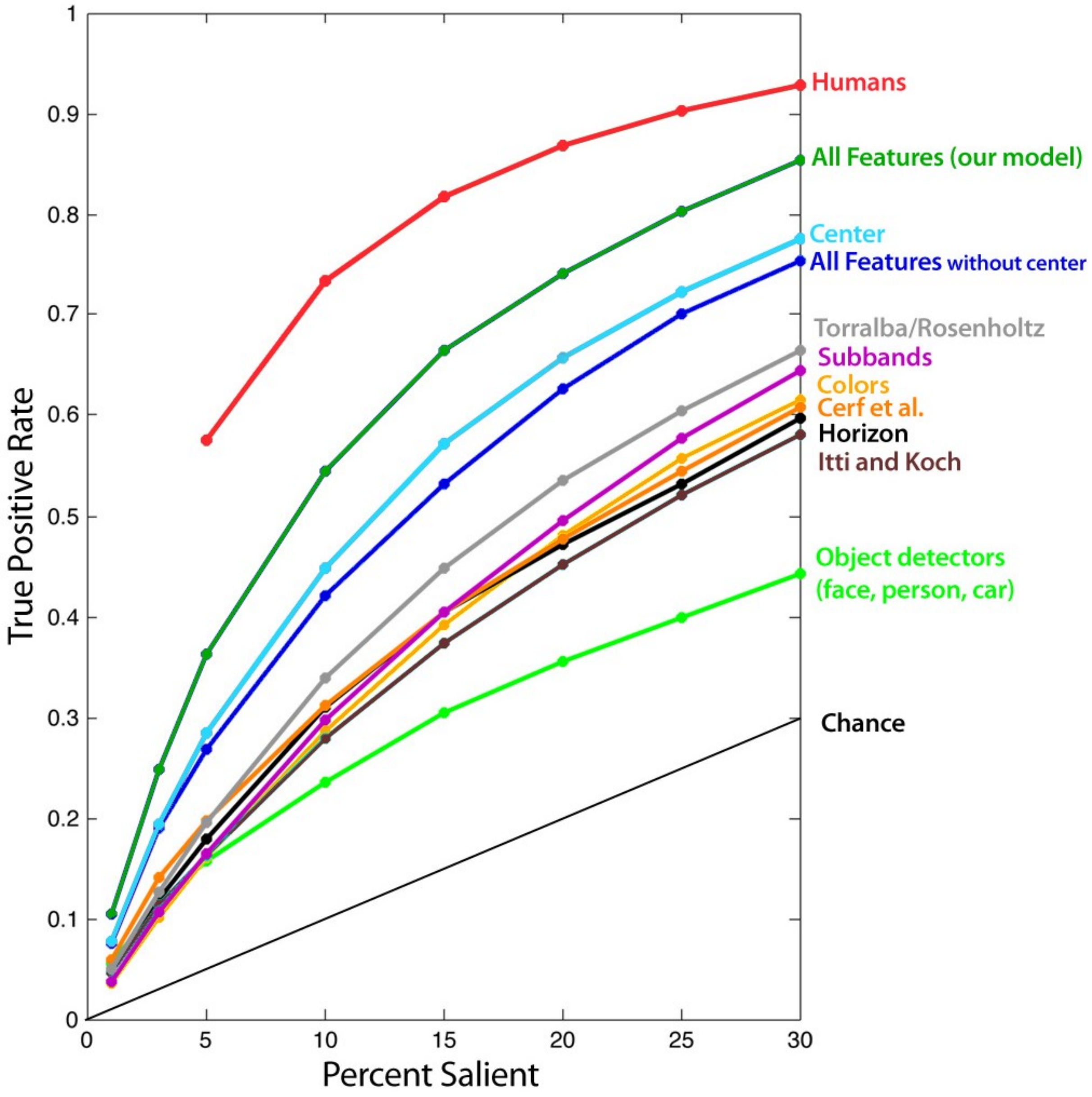
AUC Types

- Type -1

  Sweeping a threshold over the saliency map and compute the ratio of human fixations above the threshold

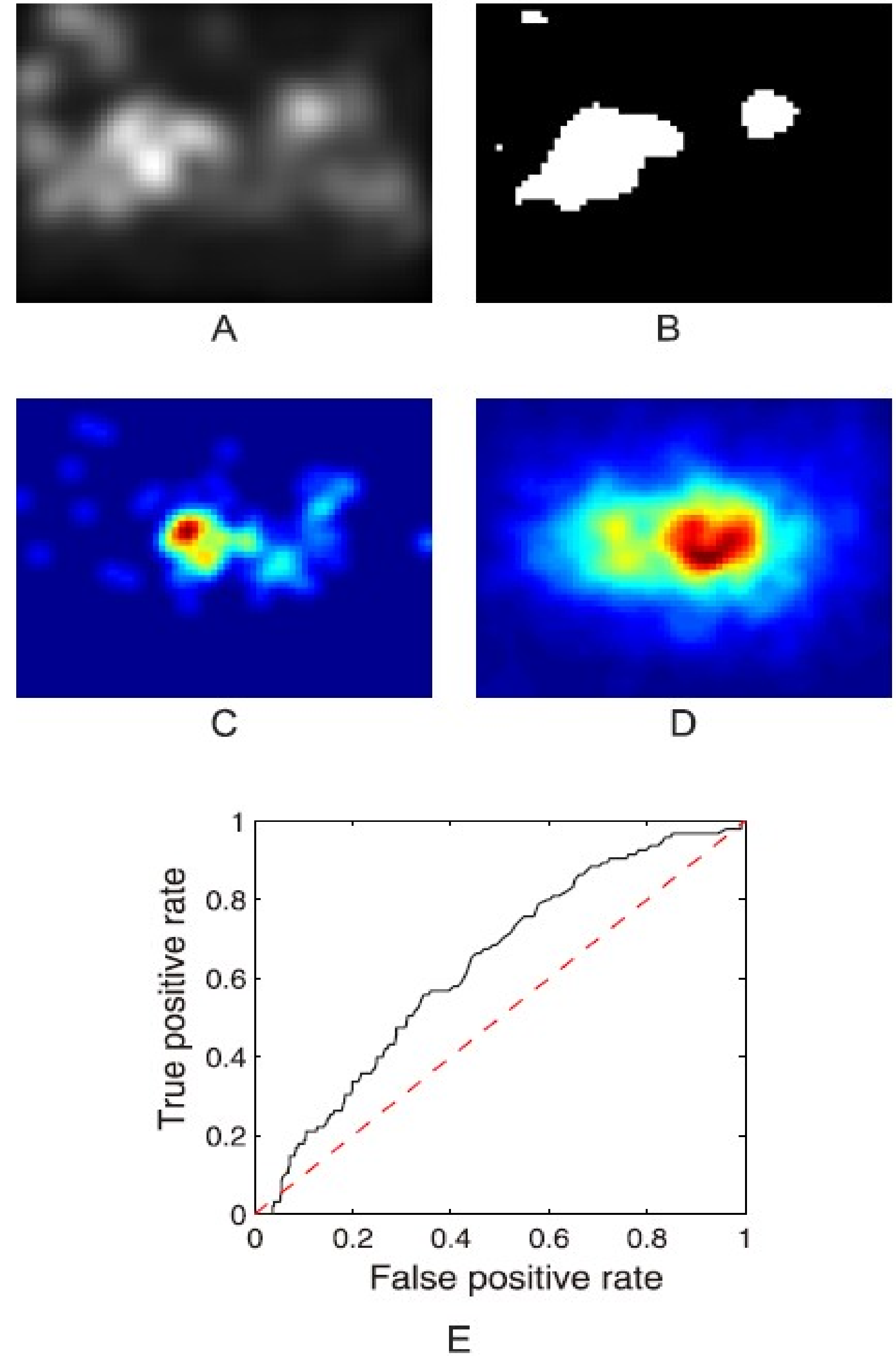
  [Judd et al. ICCV 2009]

  [Ehinger et al., 2009]



AUC Types

- Type-2
  - **Positive** samples come from the human fixations
  - **Negative** samples are uniformly generated
- Type-3
  - **Negative** samples are taken from distributions of human fixations over other images plus other subjects over the image under test
- Type-4
  - **Negative** samples are saliency values at human fixations from saliency map of another image



Adapted from Hou et al., PAMI 2012

true positive (TP)
eqv. with hit

true negative (TN)
eqv. with correct rejection

false positive (FP)
eqv. with false alarm, Type I error

false negative (FN)
eqv. with miss, Type II error

sensitivity or true positive rate (TPR)
eqv. with hit rate, recall

$$TPR = TP/P = TP/(TP + FN)$$

false positive rate (FPR)
eqv. with fall-out

$$FPR = FP/N = FP/(FP + TN)$$

accuracy (ACC)

$$ACC = (TP + TN)/(P + N)$$

specificity (SPC) or True Negative Rate

$$SPC = TN/N = TN/(FP + TN) = 1 - FPR$$

positive predictive value (PPV)
eqv. with precision

$$PPV = TP/(TP + FP)$$

negative predictive value (NPV)

$$NPV = TN/(TN + FN)$$

false discovery rate (FDR)

$$FDR = FP/(FP + TP)$$

Matthews correlation coefficient (MCC)

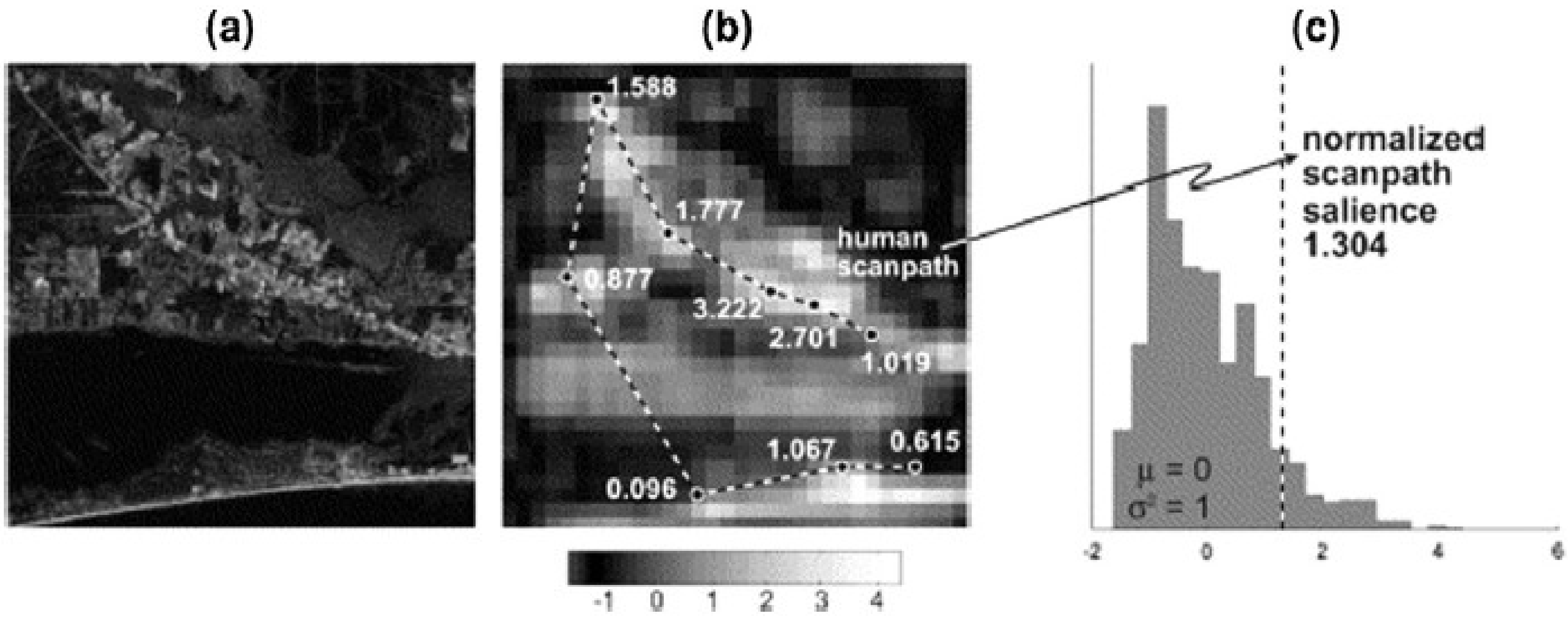
$$MCC = (TP \times TN - FP \times FN)/\sqrt{PNP'N'}$$

F1 score
is the harmonic mean of precision and recall

$$F1 = 2TP/(P + P') = 2TP/(2TP + FP + FN)$$

Source: Fawcett (2006).

Normalized Scanpath Saliency (NSS)

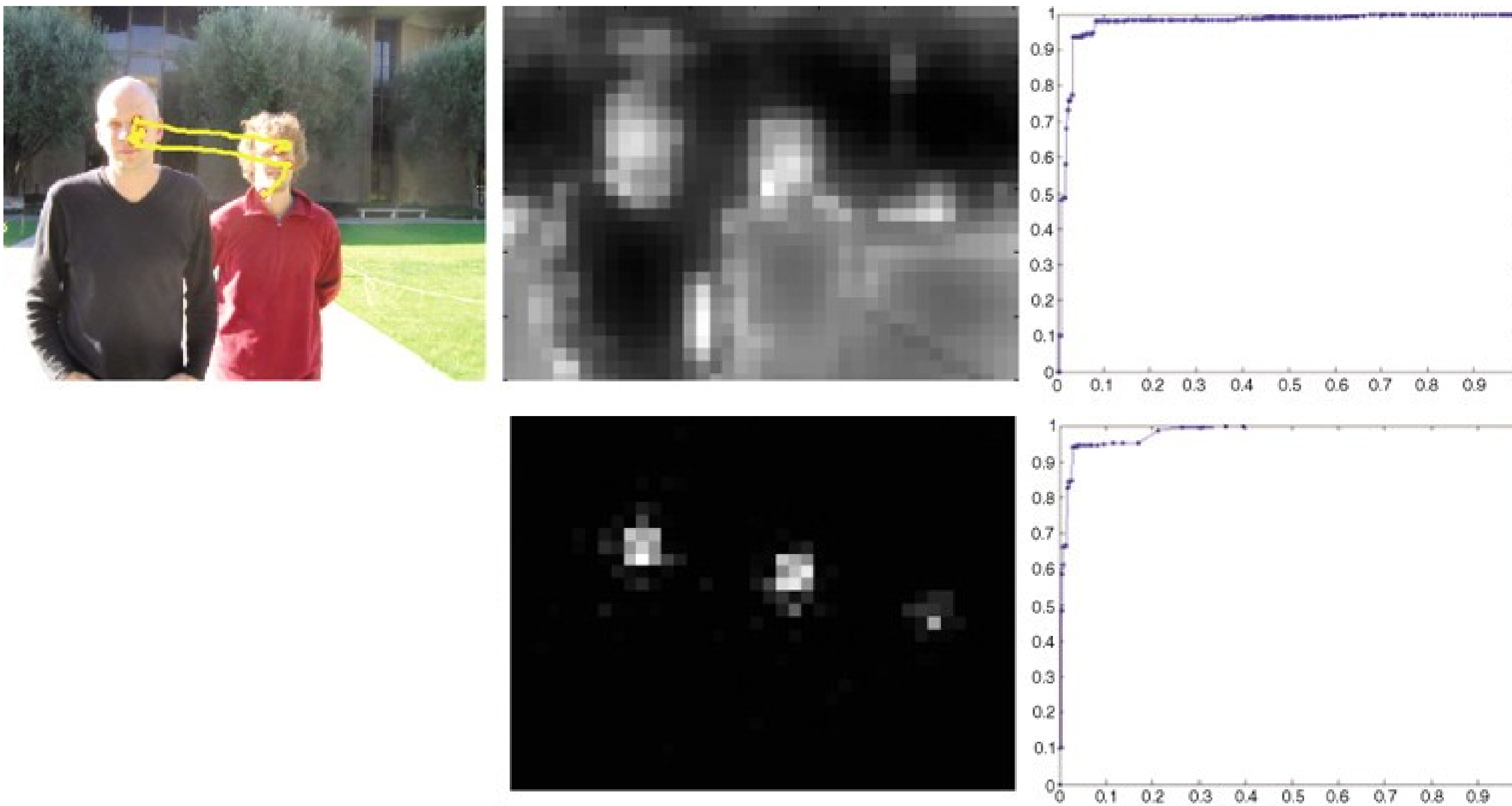


$$NSS = \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{1}{\sigma_S} \left(\overline{S_i} - \mu_S \right)$$

Peters et al., Vision Research 2005

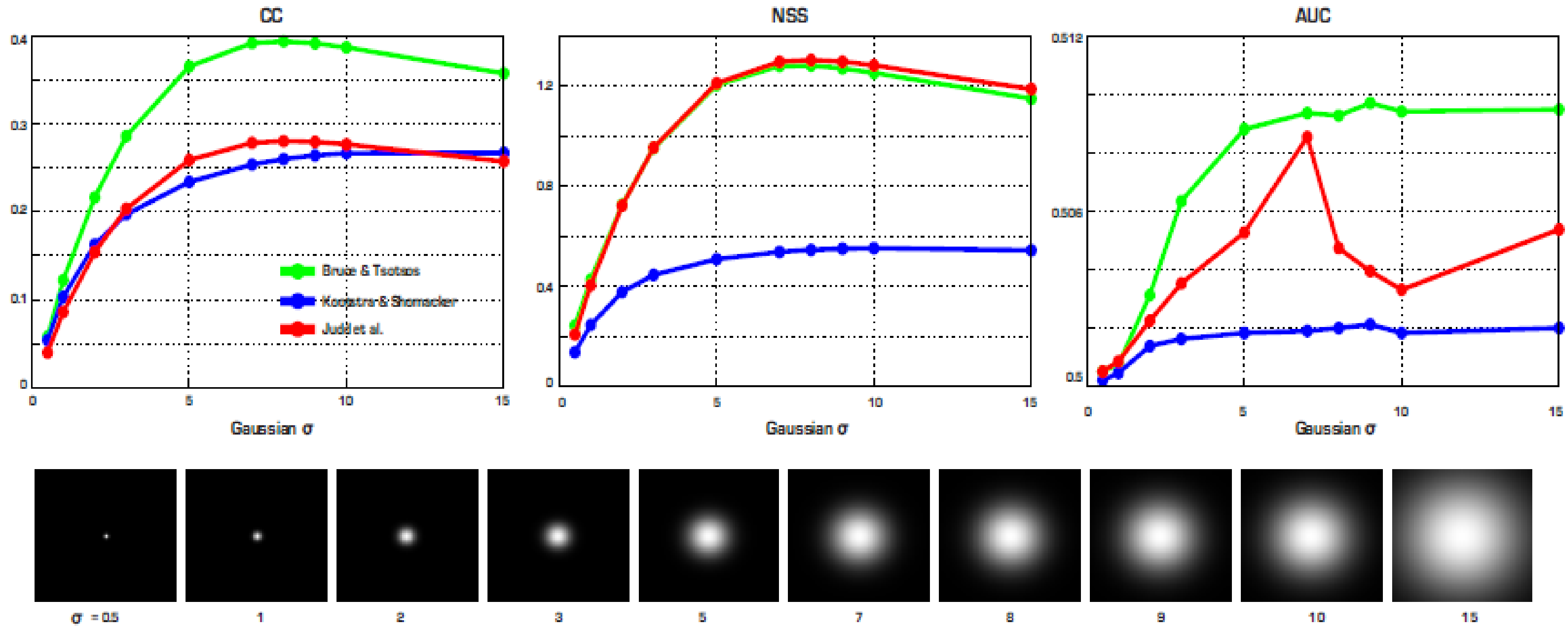
AUC ?

- As long as the hit rates are high, the AUC is always high regardless of the false alarm rate.



Zhao and Koch JOV 2011

Gaussian Blob



Plan

- Models
- Challenges
  - Dataset
  - Scores
  - Center-Bias (CB)
- Benchmark
- Summary and Future

Ali Borji

SaliencyEvaluation

- Mission
- Synthetic Patterns
- ImageDatasets
- VideoDatasets
- Evaluation measures
- PreliminaryResults
- Links
- Dicussions
- References
- Sitemap

Mission

Important note: As the very first attempt to compare saliency models and build a benchmark, we started this project in March 2010. Our paper containing results is under press in IEEE Trans. Image Processing. Please download it from my homepage.

Modeling visual attention, specially bottom-up and image-driven saliency, has been the subject of many research efforts in the past 20 years. There are many models available now which have been evaluated over different datasets using various evaluation measures.

Our mission, here is to unify the research in visual attention modeling by sharing evaluation softwares and benchmark datasets. In this direction, we have already ran and evaluated nearly 30 saliency models over synthetic images, eye movement datasets on still images and videos. Please note that, results shown here are preliminary and are subject to change.

We hope that, our efforts here helps setting some standard benchmark datasets and evaluation scores for fair evaluation of models and therefore boosting advancement in saliency modeling research.

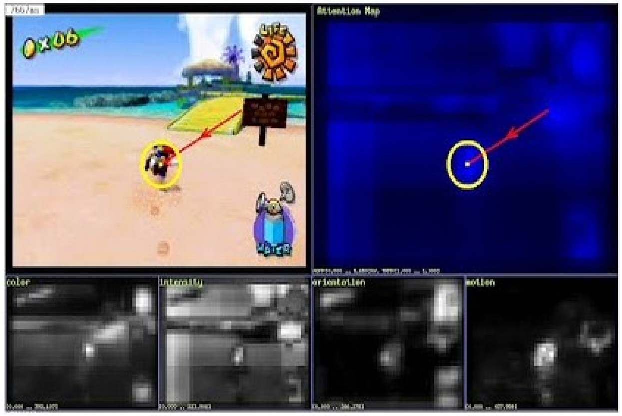
Clearly, the success of this project is highly dependent on contributions of all researchers in this field.

Note: Our results are still under review. We will re-format this website after the review process. Thanks for your patience.

Ali Borji and Laurent Itti
{borji,itti}@usc.edu

My homepage: http://ilab.usc.edu/~borji/index.html

http://ilab.usc.edu



Neuromorphic Vision C++ Toolkit (iNVT) developed at iLab, USC, http://ilab.usc.edu/toolkit/. A saccade is targetted to the location that is different from its surroundings in several feature channels. In this frame of a video, attention is strongly driven by motion saliency.

people.csail.mit.edu/tjudd/SaliencyBenchmark/

saliency benchmark

Which model of saliency best predicts where people look?

Many computational models of visual attention have been created from a wide variety of different approaches to predict where people look in images. Each model is usually introduced by demonstrating performances on new images, and it is hard to make immediate comparisons between models. To alleviate this problem, we propose a **benchmark data set containing 300 natural images with eye tracking data from 39 observers to compare model performances**. This is the largest data set with so many viewers per image. We calculate the performance of many models at predicting ground truth fixations using three different metrics: a receiver operating characteristic, a similarity metric, and the Earth Mover's Distance. We post the results here and provide a way for people to submit new models for evaluation.

paper

This benchmark is released in conjunction to the paper **"A Benchmark of Computational Models of Saliency to Predict Human Fixations"** by Tilke Judd, Fredo Durand and Antonio Torralba, available as a Jan 2012 MIT tech report.

images

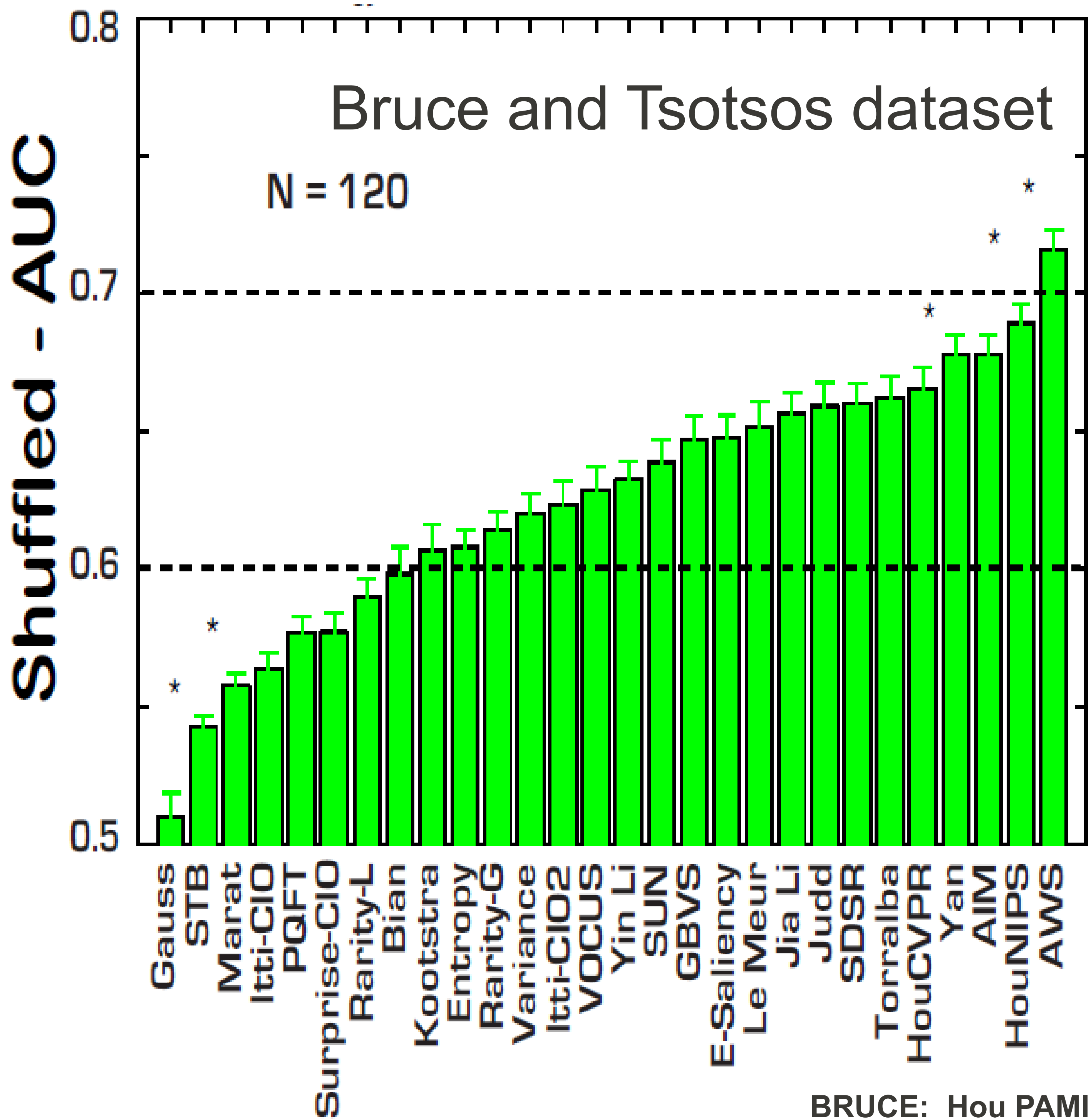
300 benckmark images (The fixations from 39 viewers per image are not public such that no model can be trained using this data set.)

A **comparison** of images and saliency maps for several models.

model performances

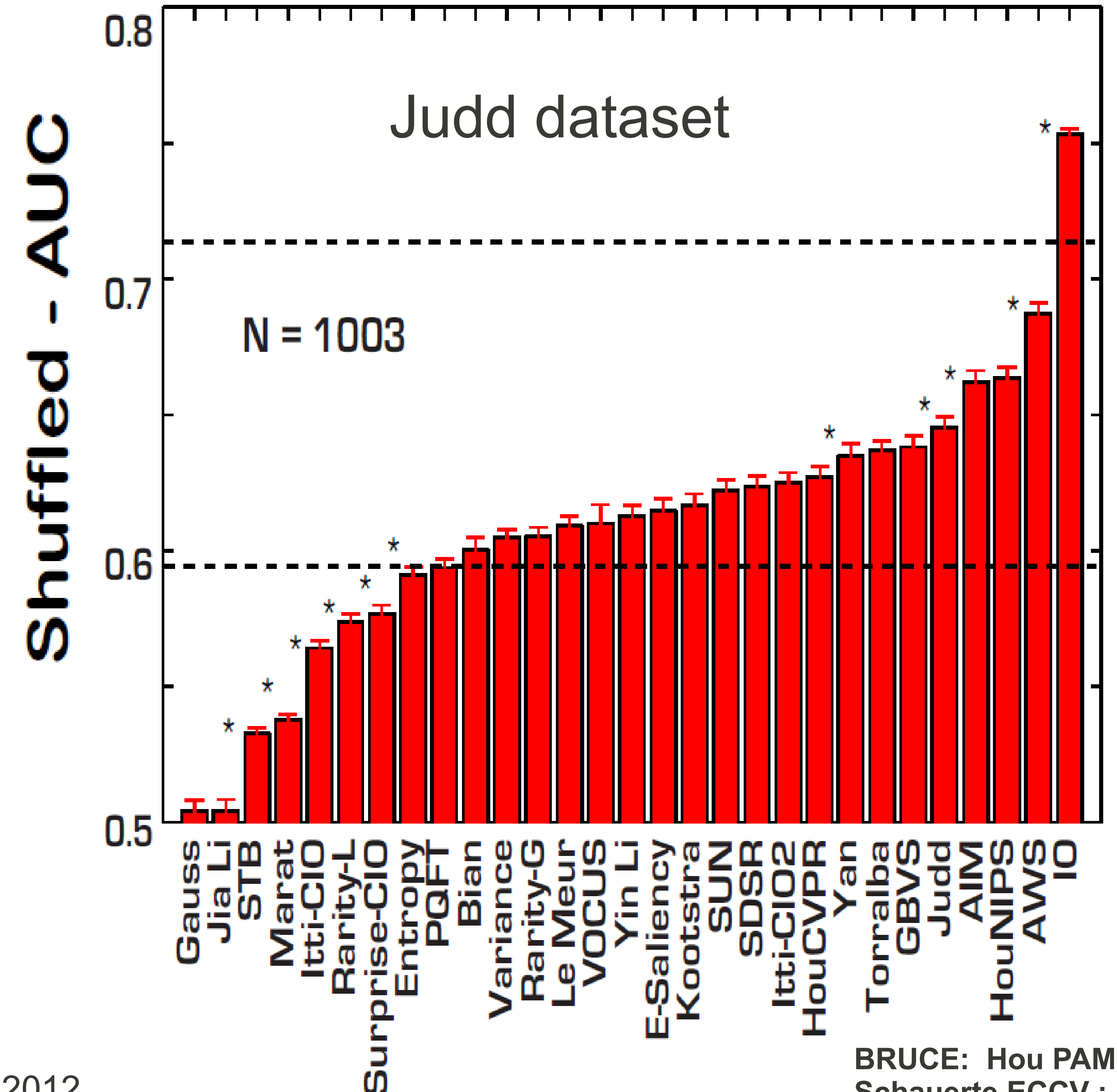
Model Name	Link to code	Area under ROC* curve (higher is better)	Similarity* (higher is better)	Earth mover's distance* (lower is better)
Humans**	code	0.922	1	0
Judd et al.	code	0.811	0.506	3.13
CovSal	paper, website	0.8056	0.5018	3.1092
Tavakoli et al. 2011	paper and website	0.8033	0.4952	3.3488
Region Contrast	website with paper	0.7922	0.4705	3.4180

Tilke Judd



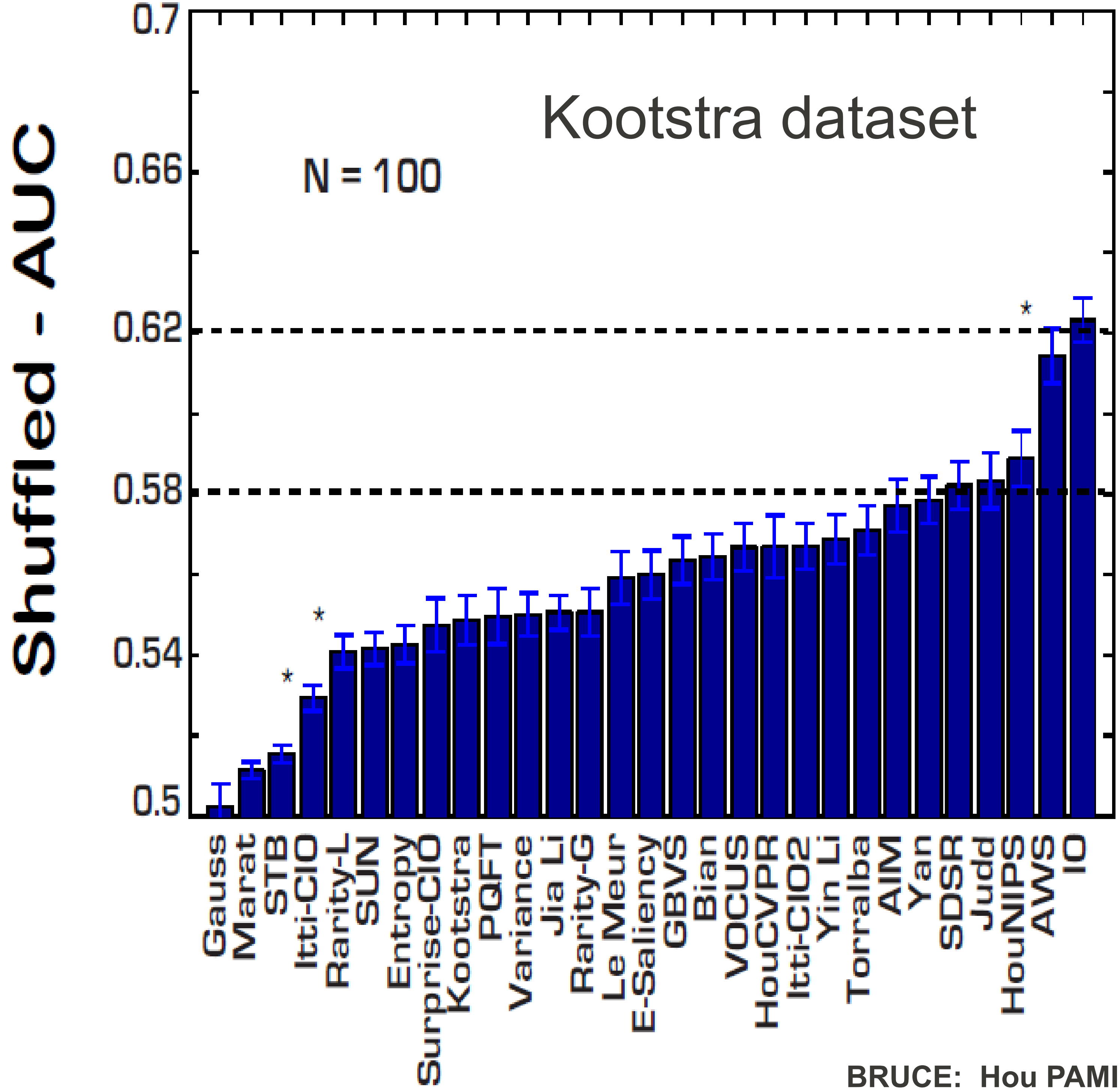
As of Sept. 2012

BRUCE: Hou PAMI: 0.7131
Schauerte ECCV : 0.7201



As of Sept. 2012

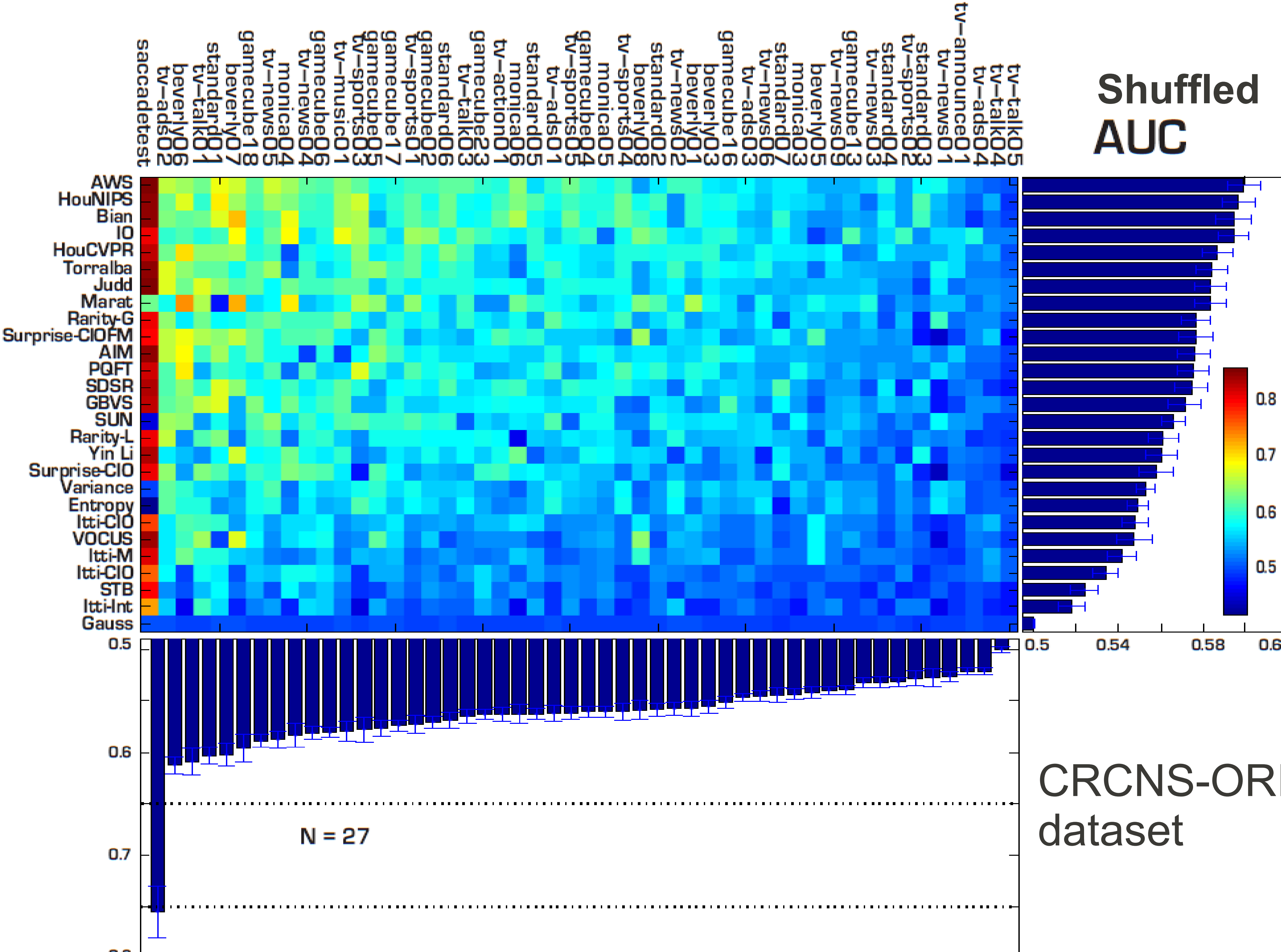
BRUCE: Hou PAMI: 0.6604
Schauerte ECCV : 0.6751



As of Sept. 2012

BRUCE: Hou PAMI: 0.6089
Schauerte ECCV : 0.6125

Shuffled AUC



CRCNS-ORIG dataset

Salient object detection models

#	Acronym (Model)	Ref.	Pub/Year	Code	Resolution	DB	Avl.
1	**IO**: Inter-observer model	-	-	M	$w \times h$	All	✓
2	**MAP**: Mean Annotation Position	-	-	M	500×500	All	✓
3	**MZ**: Ma and Zhang	[51]	ACM-M/2003	S	$w \times h$	ASD	✓
4	**LC**: Zhai and Shah	[18]	ACM-M/2006	C	$w \times h$	All	✓
5	**salLiu**: Liu *et al.*	[33]	CVPR/2007	M	max 200	All	✓
6	**AC**: Achanta *et al.*	[14]	ICVS/2008	M	$w \times h$	All	✓
7	**MSSS**: Achanta and Sušstrunk	[55]	ICIP/2009	M	$w \times h$	All	✓
8	**FTS**: Achanta *et al.*	[16]	CVPR/2009	M	$w \times h$	All	✓
9	**EDS**: Rosin	[19]	PR/2009	C	$w \times h$	All	✓
10	Gopalakrishnan *et al.*	[34]	CVPR/2009	-	-	-	-
11	Marchesotti *et al.*	[35]	ICCV/2009	-	-	-	-
12	**Valenti**: Valenti *et al.*	[40]	ICCV/2009	S	$w \times h$	ASD/MSRA	✓
13	**Goferman**: Goferman *et al.*	[15]	CVPR/2010	M	max 250	All	✓
14	**PMehrani**: Mehrani and Veksler	[23]	BMVC/2010	S	$w \times h$	ASD/SED1	✓
15	Rahtu *et al.*	[29]	ECCV/2010	-	-	-	-
16	Khuwuthyakorn *et al.*	[28]	ECCV/2010	-	-	-	-
17	Zhang *et al.*	[21]	IEEE TOM/2010	-	-	-	-
18	**JiaLiSal**: Jia Li *et al.*	[36]	IJCV/2010	S	$[w\ h]/16$	ASD/MSRA	✓
19	**LiuICIP**: Liu *et al.*	[53]	ICIP/2010	S	$w \times h$	ASD	✓
20	**MichalGazit**: Gazit *et al.*	[37]	ECCV-W/2010	M	$w \times h$	All	✓
21	**DAKlein**: Klein and Frintrop	[25]	ICCV/2011	S	$w \times h$	All	✓
22	**MengW**: M. Wang *et al.*	[18]	CVPR/2011	S	$w \times h$	ASD	✓
23	Feng *et al.*	[22]	ICCV/2011	-	-	-	-
24	Deng and Luo	[39]	OE/2011	-	-	-	-
25	Lu *et al.*	[24]	ICCV/2011	-	-	-	-
26	L. Wang *et al.*	[26]	ICCV/2011	-	-	-	-
27	**SVO**: Chang *et al.*	[27]	ICCV/2011	M	$w \times h$	All	✓
28	**CBsal**: Jiang *et al.*	[31]	BMVC/2011	M	$w \times h$	All	✓
29	**RC**: M.M. Cheng *et al.*	[13]	CVPR/2011	C	$w \times h$	All	✓
30	**HC**: M.M. Cheng *et al.*	[13]	CVPR/2011	C	$w \times h$	All	✓
31	**Materias**: Li *et al.*	[36]	BMVC/2011	M	$w \times h$	All	✓
32	**LiuIETIP**: Liu *et al.*	[42]	IEEE TIP/2011	S	$w \times h$	ASD	✓
33	**Mishra**: Mishra *et al.*	[49]	PAMI/2011	C	$w \times h$	All	✓
34	**SRS1**: Siagian and Koch	[50]	Submitted.	C	$w \times h$	All	✓

Model Ranking

As of Sept. 2012

#	Salient object detection models					Fixation prediction models				
	ASD	**MSRA**	**SED1**	**SED2**	**SOD**	**ASD**	**MSRA**	**SED1**	**SED2**	**SOD**
1	CBsal	CBsal	Gof.	RC	SVO	GBVS	GBVS	AIM	AWS	GBVS
2	LiuICIP	SVO	SVO	Gof.	Gof.	HouNIPS	HouNIPS	GBVS	GBVS	MAP
3	SVO	Gof.	CBsal	HC	MAP	AIM	AIM	MAP	SEO	AIM
4	LiuIETIP	RC	PMehrani	SVO	RC	AWS	MAP	HouNIPS	AIM	HouNIPS

SVO: Chang, K.Y., Liu, T.L., Chen, H.T., Lai, S.H.: ICCV (2011)
CBSal: Jiang, H., Wang, J., Yuan, Z., Liu, T., Zheng, N., Li, S.: BMVC (2011)
Gof.: Goferman, S., Zelnik-Manor, L., Tal, A.: CVPR (2010)
RC: Cheng, M.M., Zhang, G.X., Mitra, N.J., Huang, X., Hu, S.M.: CVPR (2011)
LiuICIP: Liu, Z., Xue, Y., Shen, L., Zhang, Z.: ICIP

Fixation prediction models score lower than salient object detection models

OPEN ACCESS Freely available online PLoS one

Measures and Limits of Models of Fixation Selection

Niklas Wilming*, Torsten Betz, Tim C. Kietzmann, Peter König

Institute of Cognitive Science, University of Osnabrück, Osnabrück, Germany

Abstract

Models of fixation selection are a central tool in the quest to understand how the human mind selects relevant information. Using this tool in the evaluation of competing claims often requires comparing different models' relative performance in predicting eye movements. However, studies use a wide variety of performance measures with markedly different properties, which makes a comparison difficult. We make three main contributions to this line of research: First we argue for a set of desirable properties, review commonly used measures, and conclude that no single measure unites all desirable properties. However the area under the ROC curve (a classification measure) and the KL-divergence (a distance measure of probability distributions) combine many desirable properties and allow a meaningful comparison of critical model performance. We give an analytical proof of the linearity of the ROC measure with respect to averaging over subjects and demonstrate an appropriate correction of entropy-based measures like KL-divergence for small sample sizes in the context of eye-tracking data. Second, we provide a lower bound and an upper bound of these measures, based on image-independent properties of fixation data and between subject consistency respectively. Based on these bounds it is possible to give a reference frame to judge the predictive power of a model of fixation selection . We provide open-source python code to compute the reference frame. Third, we show that the upper, between subject consistency bound holds only for models that predict averages of subject populations. Departing from this we show that incorporating subject-specific viewing behavior can generate predictions which surpass that upper bound. Taken together, these findings lay out the required information that allow a well-founded judgment of the quality of any model of fixation selection and should therefore be reported when a new model is introduced.

Citation: Wilming N, Betz T, Kietzmann TC, König P (2011) Measures and Limits of Models of Fixation Selection. PLoS ONE 6(9): e24038. doi:10.1371/journal.pone.0024038

Editor: Thomas Wennekers, The University of Plymouth, United Kingdom

Received February 28, 2011; **Accepted** July 29, 2011; **Published** September 12, 2011

Funding: This work was funded by the 7th Framework Programe, Grant Agreement No 270212, Extending Sensorimotor Contingencies to Cognition. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors wish to declare, for the avoidance of any misunderstanding, the following competing interests: Niklas Wilming, Torsten Betz, and Peter König hold stock in WhiteMatter Labs GmbH, who markets and sells predictions of a visual attention model. There are no patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLoS ONE policies on sharing data and materials. The authors furthermore believe that the reported results do not influence the fortune of the company (positively or negatively). Tim Kietzmann declares no competing interests.

* E-mail: nwilming@uos.de

Plan

- Models
- Challenges
  - Datasets
  - Scores
  - Center-Bias (CB)
- Benchmark
- Summary and Future

Summary

- We are much more clear than 10 years ago in saliency modeling
  - Center-bias
  - Scores
  - Datasets
  - Smoothing
- Still the landscape is not much clear in some areas
  - Center-bias
  - Top-down factors and interaction with BU factors (particularly in the spatio-temporal domain)
    - Emotions, memory, experience, etc.
  - The gap between models and humans

Open challenges

- What is the best score?
- What is the best way to tackle center-bias?
- What is the most representative dataset?
- What is the best way to take advantage of object information?
- How far are we from human performance and where?
- How to deal with model parameters?
- How to take advantage of top-down influences?
- Is there a model that works consistently the best in all test cases? natural scenes vs. psychological patterns? If not, why?
- Really how important and useful is saliency modeling for hard problems in computer vision? e.g., object recognition?

Applications

- Marketing and Advertisement
- Intent decoding / Mind reading
  - Patient vs. normal observer
- computer vision
  - Image re-targeting, segmentation,
  - compression, detection, recognition,
  - enhancement, etc.
- Robotics
  - Localization and Navigation
  - Human robot interaction, etc.

Future directions

- **Top-down attention**
  - Visual Search and Object Detection
  - Global Context
  - Real-world visio-motor
    - Ego centric and 1st person vision
  - 3D/Social scenarios (Sheikh et al. NIPS 2012)
  - Gaze and pointing direction
  - Spatio-temporal (actor, action, etc.)
  - Interactive environments and different tasks
  - Sequential attention for object recognition
  - Action recognition
- **Model Comparison and Benchmarking**
  - **Evaluation measures**
    - Several scores ?!
    - Scanpath (e.g., Le meur)
  - **Datasets**
    - Emotions and saliency (Subramanian et al., NUSEF dataset)
  - **Model comparison in applications**

- **Applications**
  - Image compression (Itti et al. TIP 2004)
  - Steganography and biometrics (i.e., observer decoding)
  - Image re-targeting
  - Image aesthetics
  - Robotics; navigation, localization, etc [e.g., Simone Frintrop]
  - Human computer/robot interaction
  - Web search [+ clicks]
  - Mind reading/ Intent decoding (e.g., using Google glass)

Behavioral labs

- Hayhoe and Ballard; UT Austin
- Tatler; Dundee
- John Henderson
- Laurent Itti
- Peter Koning
- Wolfgang Einhauser
- Foulsham
- Susana Martinez Condes
- Jeremy Wolfe; Hardvard
- Marc Pomplun; Boston Uni.
- Maria Carasco
- Nakayama
- John Tsotsos
- Aude Oliva & Torralba
- Gregory Zelinsky
- Castelhano
- Cave
- Kastner
- Pelz
- John Reynoldz
- Chellazi
- Grossberg
- Deco
- Rolls

- Matt Peterson
- Edward Awh
- Edward K. Vogel
- Tirin Moore
- Bisley
- Posner
- Yantis
- Kanwisher
- Desimone
- Jiang lab at University of Minnesota,
- Jochen Triesch
- Ronald Rensink
- ROGER REMINGTON
- Lester loshkey
- Stephen Mitroff
- Yesherun
- Christof Koch
- Hannes Schulz
- Andreas K. Engel
- Milanfar
- Ueli Rutishauser
- Andreas K. Engel

- Preeti Vergeese
- Michael land
- Rajesh Rao
- Munz
- Rothkopf
- Martin Eimer
- Garry Cottreel
- Movellan
- Geisler
- Michael Mozer
- Nuno Vasconcelos
- Wolfgang Einhauser
- Antje Nuthman
- Geisler
- Thorpe
- Andrew Hollingworth
- Monica Castelhano
- Andrew Hollingworth
- David Heeger

You can meet some of these people in VSS of SFN conferences.

SfN Society for Neuroscience

Advancing the Understanding of the Brain and Nervous System

Vision Sciences Society

understanding vision and brain

Neuroscience
Psychophysics
Cognition
Action
Computation
Development

Some remarks

- Scoring and model comparison is a serious issue
- Papers are often cited inappropriately [specially in behavioral venues such as JOV]
    - Some used shuffled AUC and some AUC (both at JOV)

Open Forum

- Please:
  - Try to ask questions and initiate discussions
  - Correct us if you think we are wrong somewhere
  - Share your comments with us and other fellow researchers
  - Let us know your critics and opinions
  - etc.