



Saliency
Detection

Background
Development
Features
Priors
Aggregation

Saliency Detection

显著性检测

Zhu Yafei

CVBIOUC

<http://vision.ouc.edu.cn/~zhenghaiyong>

March 24, 2015



Contents

Saliency
Detection

Background
Development
Features
Priors
Aggregation

1 Background

2 Development

3 Features

4 Priors

5 Aggregation



What?

Saliency
Detection

Background
Development
Features
Priors
Aggregation

Saliency(显著)

■ 明显、引入注目



What?

Saliency
Detection

Background
Development
Features
Priors
Aggregation

Saliency(显著) \iff
■ 明显、引入注目



What?

Saliency
Detection

Background
Development
Features
Priors
Aggregation

Saliency(显著) \Leftarrow Visual
■ 明显、引入注目 Attention



What?

Saliency
Detection

Background

Development

Features

Priors

Aggregation

A rich stream of visual data enters our eyes every second.¹

¹K. Koch and J. McLean, “How much the eye tells the brain”, Current Biology, 2006.



What?

Saliency
Detection

Background

Development

Features

Priors

Aggregation

A rich stream of visual data enters our eyes every second.¹

$$10^8 \sim 10^9 \text{ bits}$$

¹K. Koch and J. McLean, “How much the eye tells the brain”, Current Biology, 2006.



What?

Saliency
Detection

Background
Development
Features
Priors
Aggregation



分辨率:
500*375



What?

Saliency
Detection

Background
Development
Features
Priors
Aggregation



分辨率:
 500*375

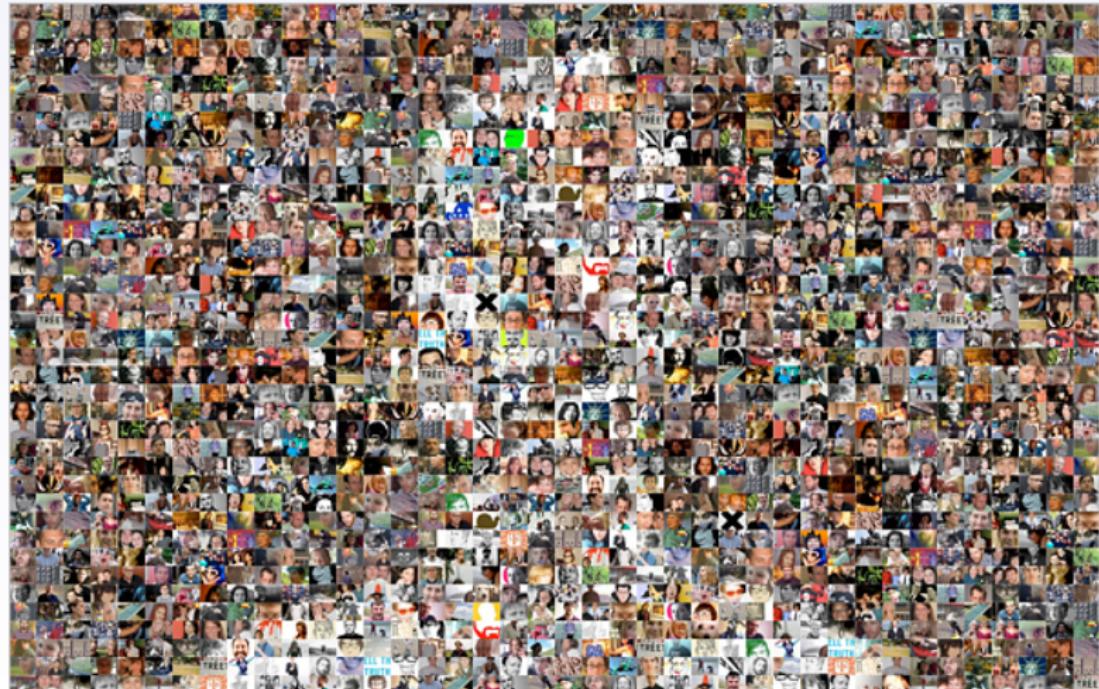
$$500 \times 375 \times 3 \times 8 = 4500000 \text{ bits}$$



What?

Saliency
Detection

Background
Development
Features
Priors
Aggregation





Vision as data reduction

Saliency
Detection

Background

Development

Features

Priors

Aggregation

- Raw feed from camera/eyes:
 - 10^{7-9} Bytes/s
- Extraction of edges and salient features
 - 10^{3-4} Bytes/s
- High-level interpretation of scene
 - 10^{1-2} Bytes/s



Visual Attention

Saliency
Detection

Background

Development

Features

Priors

Aggregation





Visual Attention

Saliency
Detection

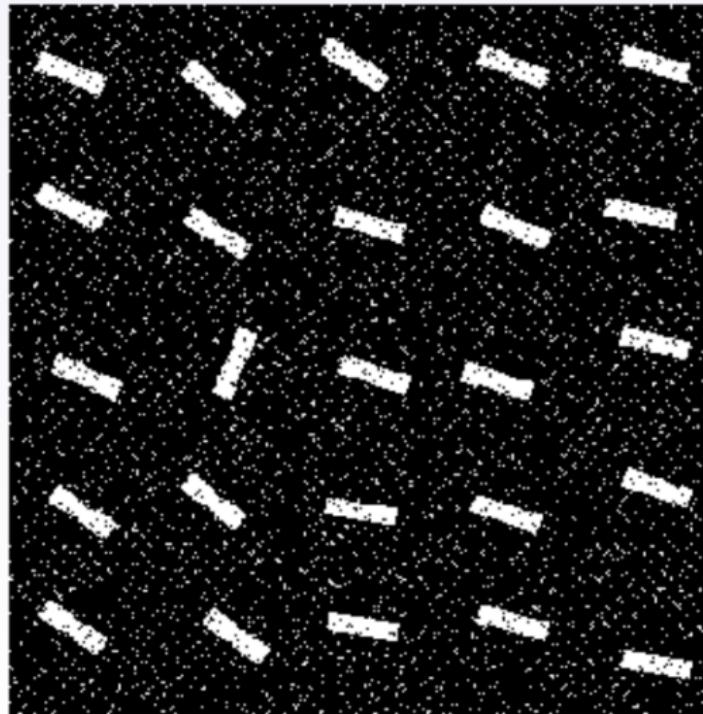
Background

Development

Features

Priors

Aggregation





Visual Attention

Saliency
Detection

Background

Development

Features

Priors

Aggregation

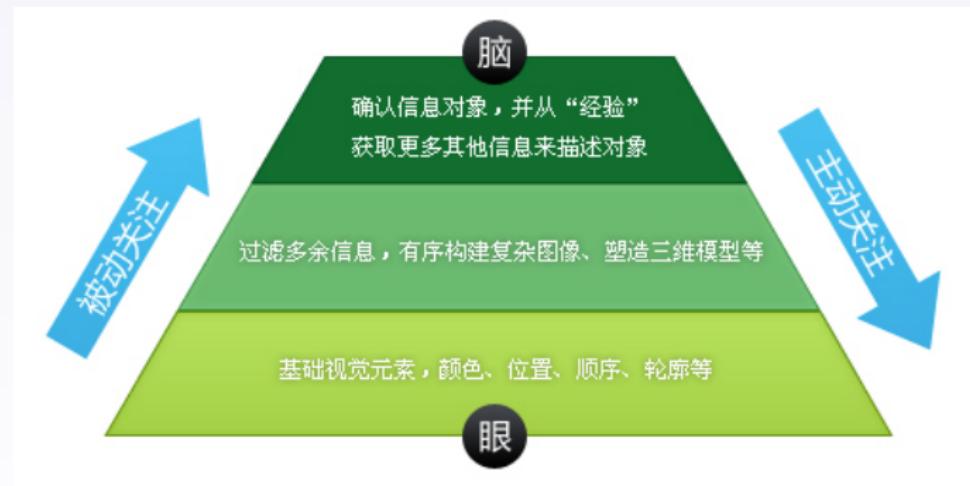




Bottom-up vs Top-down

Saliency
Detection

Background
Development
Features
Priors
Aggregation

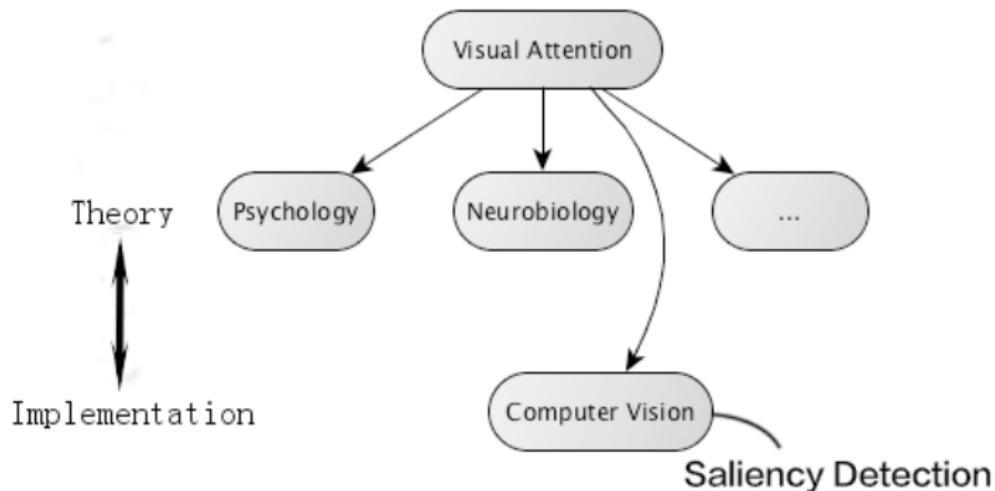




Visual Attention

Saliency
Detection

Background
Development
Features
Priors
Aggregation





Theory

Saliency
Detection

Background

Development

Features

Priors

Aggregation

- “*Feature Integration Theory*”, Treisman & Gelade, 1980.
- “*Guided Search Theory*”, Wolfe & Cave, 1989.
- “*Integrated Competition Theory*”, Desimone & Duncan, 1995.



Framework

Saliency
Detection

Background
Development
Features
Priors
Aggregation

- Koch and Ullman² proposed a feed-forward model to combine features and introduced the concept of a *saliency map* which is a topographic map that represents conspicuousness of scene locations.
- Zoom Lens Model³

²C. Koch and S. Ullman, “Shifts in selective visual attention: towards the underlying neural circuitry”, Human Neurobiology, 1985.

³Eriksen, C.W., J.D. St. James. “Visual attention within and around the field of focal attention: a zoom lens model”, Perception and Psychophysics, 1986.

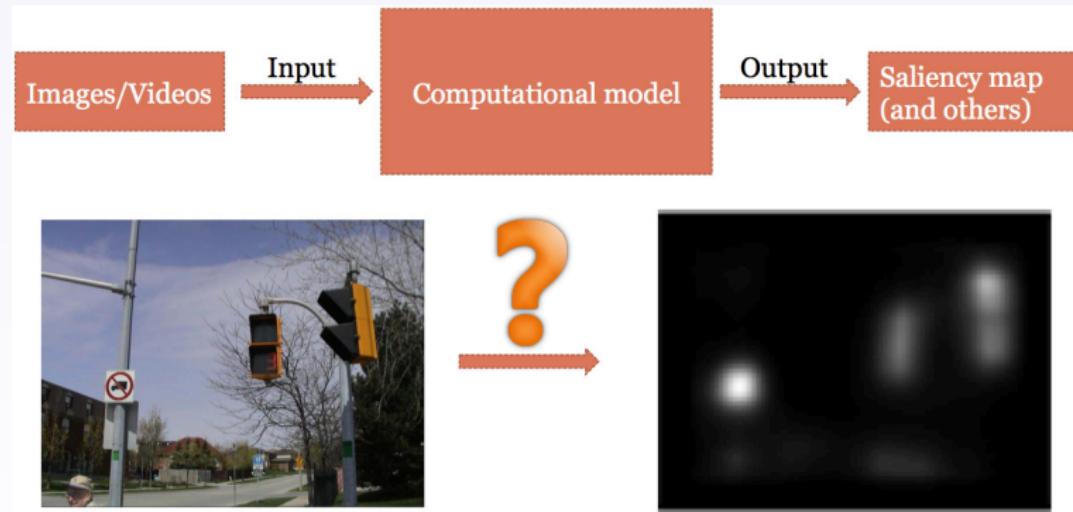


Basic Structure of Computational Models

Saliency
Detection

Background
Development
Features
Priors

Aggregation

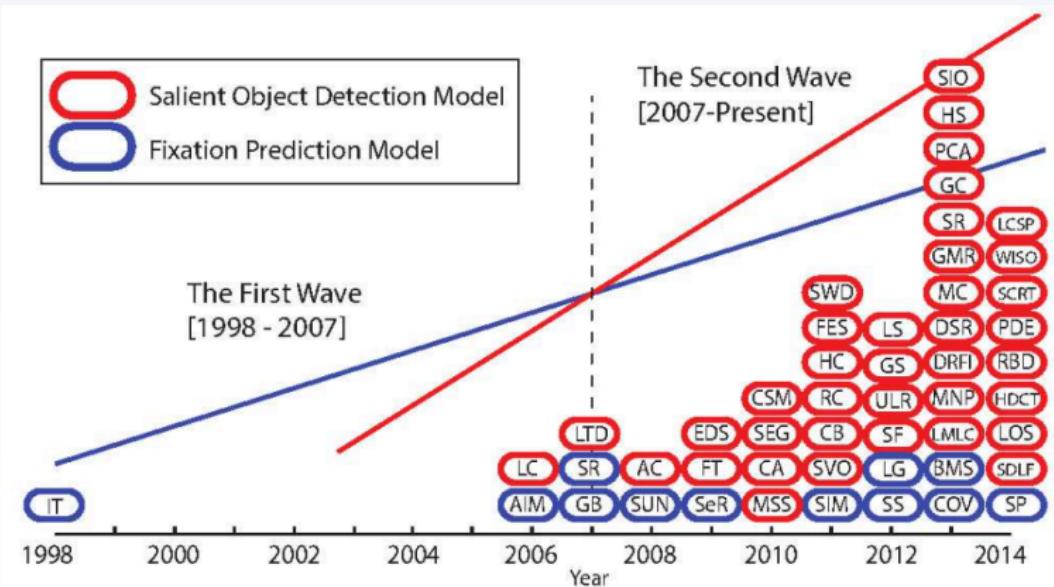




Two Waves⁴

Saliency
Detection

Background
Development
Features
Priors
Aggregation



⁴Ali Borji, Ming-Ming Cheng, "Salient object detection: a survey", 2014.



Fixation Prediction

Saliency
Detection

Background
Development

Features

Priors

Aggregation

Fixation prediction models are constructed originally to understand human visual attention and eye movement prediction.





Fixation Prediction: example

Saliency
Detection

Background
Development
Features
Priors
Aggregation





Fixation Prediction: example

Saliency
Detection

Background
Development
Features
Priors
Aggregation





Fixation Prediction: example

Saliency
Detection

Background
Development
Features
Priors
Aggregation





Fixation Prediction: example

Saliency
Detection

Background
Development
Features
Priors
Aggregation





Fixation Prediction: example

Saliency
Detection

Background
Development
Features
Priors
Aggregation





Fixation Prediction: example

Saliency
Detection

Background
Development
Features
Priors
Aggregation





Representative Work

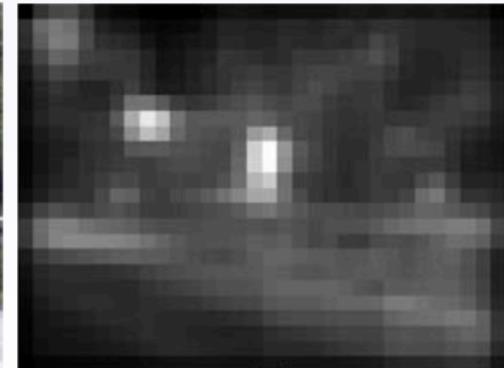
Saliency
Detection

Background
Development
Features

Priors

Aggregation

- A model of saliency-based visual attention for rapid scene analysis. PAMI 1998, Itti et al.

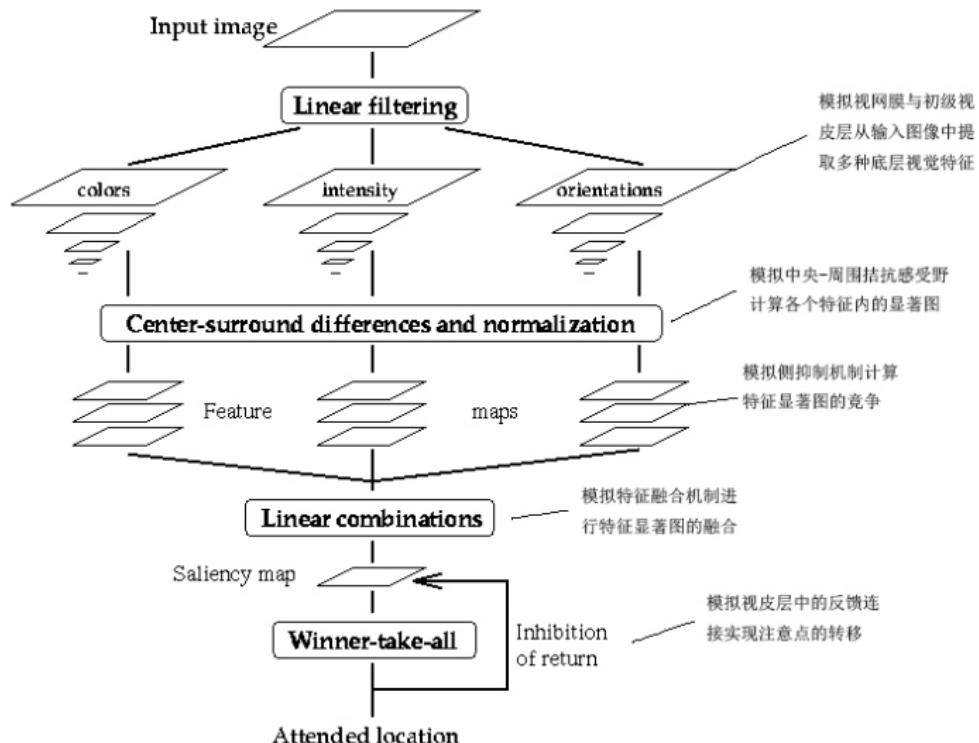




Architecture

Saliency
Detection

Background
Development
Features
Priors
Aggregation

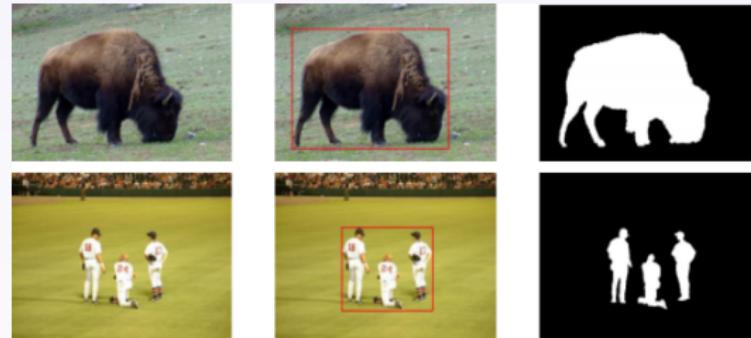




Salient Object Detection

Saliency
Detection

Background
Development
Features
Priors
Aggregation

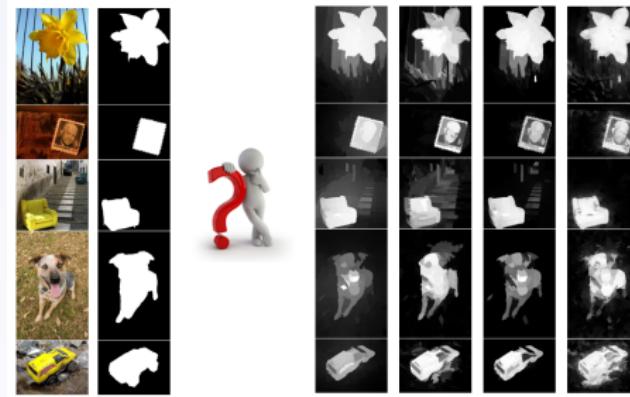


- Learning to detect a salient object. CVPR 2007, Tie Liu et al.
- Frequency-tuned salient region detection. CVPR 2009, Achanta et al.



Latest research achievements

Saliency
Detection



Background
Development
Features
Priors
Aggregation

From left to right: PBS⁵、DRFI⁶、RBD⁷、HDCT⁸

⁵Chuan Yang et al., “Graph-regularized saliency detection with convex-hull-based center prior, IEEE Signal Processing Letters, 2013.

⁶Huaizu Jiang et al., “Salient object detection: a discriminative regional feature integration approach”, in CVPR, 2013.

⁷Wangjiang Zhu et al., “Saliency optimization from robust background detection”, in CVPR, 2014.

⁸Jiwhan Kim et al., “Salient region detection via high-dimensional color transform”, in CVPR, 2014.



Why?

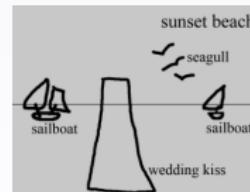
Saliency
Detection

Background
Development
Features
Priors
Aggregation

The emergence of salient object detection models is driven by the requirement of saliency-based applications.



Content aware resizing



Object manipulation



Image montage

Image collage



What can saliency not do?

Saliency
Detection

Background
Development
Features
Priors
Aggregation

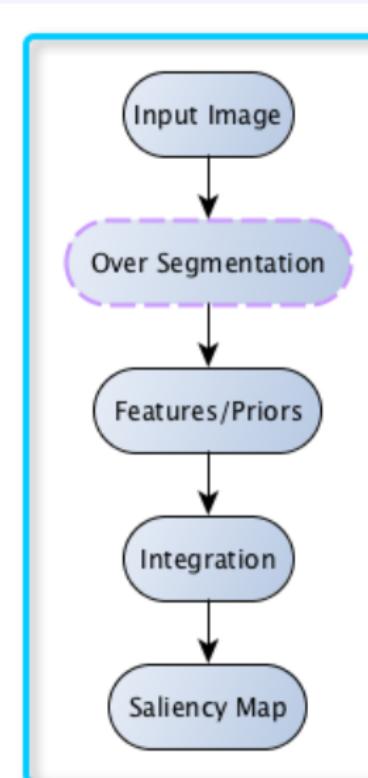




General Procedure

Saliency
Detection

Background
Development
Features
Priors
Aggregation





Features

Saliency
Detection

Background
Development

Features

Priors

Aggregation

Salient object detection: A discriminative regional feature integration approach, CVPR 2013

■ 26 维颜色纹理等对比度特征

Color and texture features		dim	Differences of features		Contrast	Backgroundness
	features		definition	dim		
a_1	the average RGB values	3	$d(a_1^R, a_1^S)$	3	$c_1 \sim c_3$	$b_1 \sim b_3$
a_2	the average L*a*b* values	3	$d(a_2^R, a_2^S)$	3	$c_4 \sim c_6$	$b_4 \sim b_6$
r	the absolute response of LM filters	15	$d(r^R, r^S)$	15	$c_7 \sim c_{21}$	$b_7 \sim b_{21}$
r	the max response among the LM filters	1	$d(r^R, r^S)$	1	c_{22}	b_{22}
h_1	the L*a*b* histogram	$8 \times 16 \times 16$	$\chi^2(h_1^R, h_1^S)$	1	c_{23}	b_{23}
h_2	the hue histogram	8	$\chi^2(h_2^R, h_2^S)$	1	c_{24}	b_{24}
h_3	the saturation histogram	8	$\chi^2(h_3^R, h_3^S)$	1	c_{25}	b_{25}
h_4	the texton histogram	65	$\chi^2(h_4^R, h_4^S)$	1	c_{26}	b_{26}

■ 34 维的区域特征

description		notation	dim	description		notation	dim
the average normalized x coordinates	p_1	1		the average normalized y coordinates	p_2	1	
the normalized perimeter	p_7	1		the 10th percentile of the normalized x coordinates	p_3	1	
the aspect ratio of the bounding box	p_8	1		the 10th percentile of the normalized y coordinates	p_4	1	
the variances of the RGB values	$p_9 \sim p_{11}$	3		the 90th percentile of the normalized x coordinates	p_5	1	
the variances of the L*a*b* values	$p_{12} \sim p_{14}$	3		the 90th percentile of the normalized y coordinates	p_6	1	
the variances of the HSV values	$p_{15} \sim p_{17}$	3		the variance of the response of the LM filters	$p_{18} \sim p_{32}$	15	
the normalized area	p_{33}	1		the normalized area of the neighbor regions	p_{34}	1	



Features

Saliency
Detection

Background
Development
Features
Priors
Aggregation

- Pixel-based
 - Contrast: $U_p(x) = \sum_{x' \in I \setminus \{x\}} D(x', x)$
 - Property: Geometric features
- Patch/Region/Superpixel-based
 - Contrast: $U_r(A_i) = \sum_{1 \leq k \leq N, k \neq i} |\Lambda| D(\Lambda_k, \Lambda_i)$
 - Property: Geometric features, Appearance features
- Multi-scale based



① Contrast

Saliency
Detection

Background

Development

Features

Priors

Aggregation

- Color
- Texture



Color

Saliency
Detection

Background

Development

Features

Priors

Aggregation

- *Color* is the visual perceptual property corresponding in humans to the categories called red, blue, yellow and others.
- *Color Space* is defined to identify colors numerically by their coordinates.



Three Elements of Color

Saliency
Detection

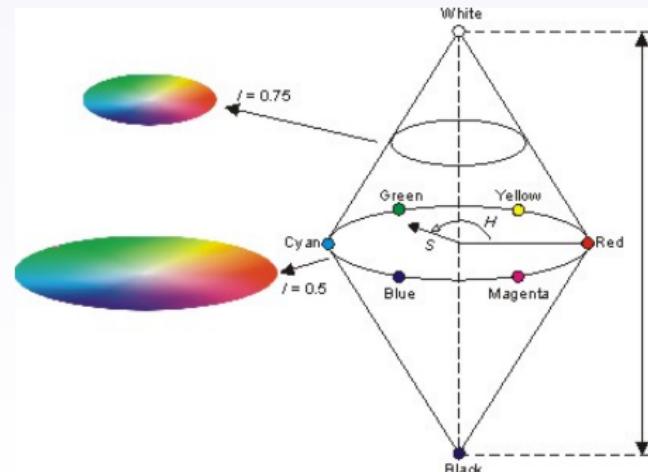
Background
Development

Features

Priors

Aggregation

- Hue: What we think as “color”—yellow, orange, cyan and magenta are examples of different hues
- Saturation: Saturation refers to the relative purity or the amount of white light mixed with a hue
- Intensity





Coding methods for humans

Saliency
Detection

Background

Development

Features

Priors

Aggregation

- **RGB** is an additive system (add colors to black) used for displays.
- **CMY** is a subtractive system for printing.
- **HSI** is a good perceptual space for art, psychology, and recognition.
- **YIQ** is used for TV and is good for compression.



Color Space

Saliency
Detection

Background

Development

Features

Priors

Aggregation

- RGB
- Lab
- HSV



Color Histogram (**Demo**)

Saliency
Detection

Background
Development
Features
Priors
Aggregation

- 对图像的颜色空间进行量化
- 计算矩阵 Q
- 超像素分割
- 对每个超像素区域，依据其中像素点的坐标值，求出对应的矩阵 Q 中的值
- 求区域之间的直方图的对比度



What is texture?

Saliency
Detection

Background

Development

Features

Priors

Aggregation

Texture is actually a very nebulous concept, often attributed to human perception, as either the feel or the appearance of (woven) fabric.⁹



⁹Mark S.Nixon and Alberto S.Aguado, “Feature Extraction & Image Processing for Computer Vision”, Publishing House of Electronics Industry, 2013.



LBP(Local Binary Pattern, 局部二值模式)

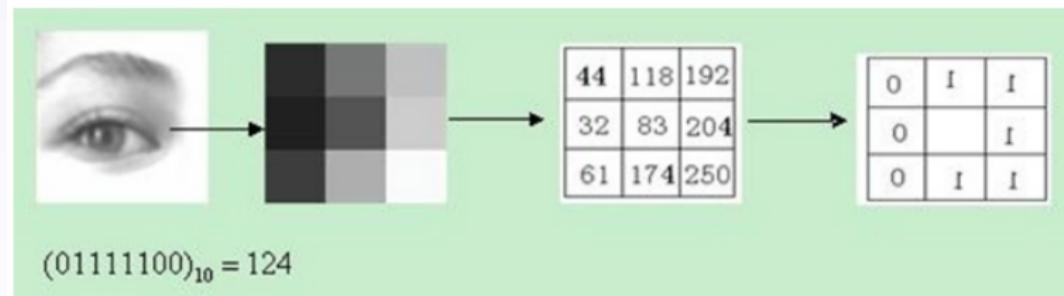
Saliency
Detection

Background
Development

Features

Priors

Aggregation





LBP

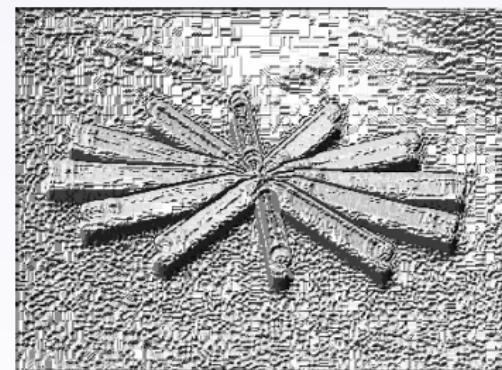
Saliency
Detection

Background
Development

Features

Priors

Aggregation





② Regional property descriptor

Saliency
Detection

Background

Development

Features

Priors

Aggregation

- Geometric feature descriptor
 - size
 - perimeter
 - position
 - area of the neighbor regions
- Appearance feature descriptor
 - the variances of the RGB values
 - the variances of the response of the LM filters



Priors

Saliency
Detection

Background

Development

Features

Priors

Aggregation

- 1 Center Prior/Location Prior
- 2 Backgroundness Prior
- 3 Boundary Connectivity Prior
- 4 Color Prior
- 5 Objectness Prior
- 6 Smoothness Prior



① Center Prior/Location Prior (**Demo**)

Saliency
Detection

Background
Development
Features

Priors
Aggregation

Objects near the image center are more attractive to people.¹⁰

This prior can be simply and effectively modeled as a Gaussian map.



$$S_D(\mathbf{x}) = \exp\left(-\frac{||\mathbf{x} - \mathbf{c}||_2^2}{\sigma_D^2}\right) \quad (1)$$

¹⁰Zhang, Lin and Gu, Zhongyi and Li, Hongyu, “SDSP: A novel saliency detection method by combining simple priors”, in ICIP, 2013.



① Center Prior/Location Prior

Saliency
Detection

Background

Development

Features

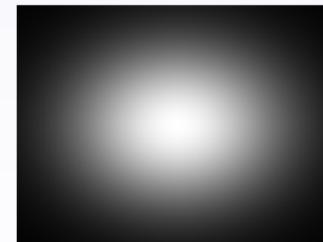
Priors

Aggregation

The salient object in an image is most probably placed near the center of the image.¹¹

Gaussian falloff weight:

$$w_i^{(n)} = \exp(-9(dx_i^{(n)})^2/w^2 - 9(dy_i^{(n)})^2/h^2) \quad (2)$$



¹¹Jiang, Huaizu and Wang, Jingdong, "Automatic salient object segmentation based on context and shape prior", in BMVC, 2011



Saliency
Detection

Background

Development

Features

Priors

Aggregation

① Center Prior/Location Prior

Assigning higher saliency to the image elements near the image center becomes invalid when the objects are placed far off the image center¹².

- Compute a convex hull enclosing interesting points to estimate the location of salient region.
- Use the centroid of the convex hull as the center to get the convex-hull-based center prior map.



¹²Yang, Chuan and Zhang, Lihe and Lu, Huchuan, "Graph-regularized saliency detection with convex-hull-based center prior", in Signal Processing Letters, IEEE, 2013



② Backgroundness Prior

Saliency
Detection

Background
Development
Features
Priors

Aggregation

Backgroundness prior is more general than center prior because salient objects can be placed off the center, but they seldom touch the image boundary.¹³

- Assuming that a narrow border of the image is background region, regional saliency can be computer as the contrast versus “background”.

¹³Wei, Yichen and Wen, Fang and Zhu, Wangjiang, “Geodesic saliency using background priors”, in ECCV, 2012.



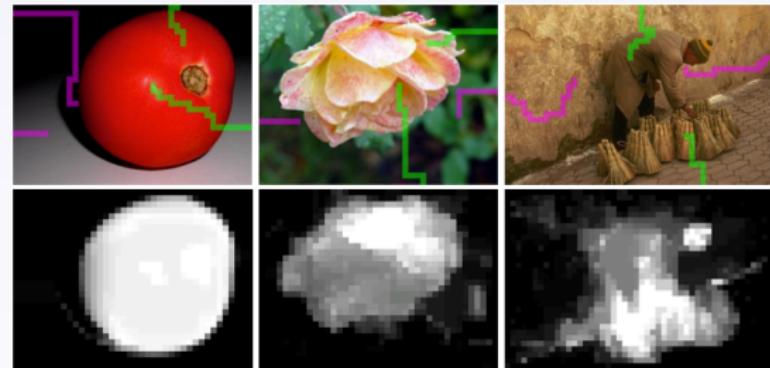
② Backgroundness Prior (**Demo**)

Saliency
Detection

Background
Development
Features

Priors

Aggregation



$$saliency(P) = \min_{P_1 = P, P_2, \dots, P_n = B} \sum_{i=1}^{n-1} weight(P_i, P_{i+1}), \quad (3)$$
$$s.t. (P_i, P_{i+1}) \in \varepsilon$$

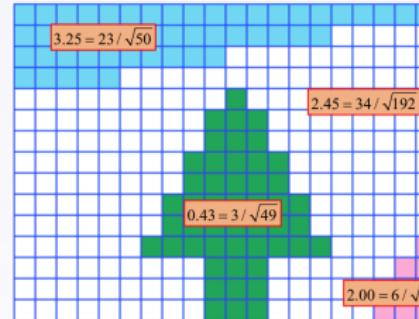


Saliency Detection

Background
Development
Features
Priors
Aggregation

③ Boundary Connectivity Prior

Object regions are much less connected to image boundaries than background ones.¹⁴



Boundary connectivity is defined to quantify how heavily a region R is connected to the image boundaries.

$$BndCon(R) = \frac{|\{p | p \in R, p \in Bnd\}|}{\sqrt{|\{p | p \in R\}|}} \quad (4)$$

¹⁴Zhu, Wangjiang and Liang, Shuang, “Saliency optimization from robust background detection”, in CVPR, 2014.



④ Color Prior (Demo)

Saliency
Detection

Background

Development

Features

Priors

Aggregation

Warm colors, such as red and yellow, are more pronounced to the human visual system than cold colors, such as green and blue.¹⁵

$$f_{an}(\mathbf{x}) = \frac{f_a(x) - mina}{maxa - mina}, f_{bn}(\mathbf{x}) = \frac{f_b(x) - minb}{maxb - minb} \quad (5)$$

$$S_c(\mathbf{x}) = 1 - exp\left(-\frac{f_{an}^2(\mathbf{x}) + f_{bn}^2(\mathbf{x})}{\sigma_c^2}\right) \quad (6)$$

- a^* -channel represents green-red information
- b^* -channel represents blue-yellow information

¹⁵Zhang, Lin and Gu, Zhongyi and Li, Hongyu, "SDSP: A novel saliency detection method by combining simple priors", in ICIP, 2013.



⑤ Objectness Prior

Saliency
Detection

Background
Development
Features
Priors
Aggregation

Objectness is defined as the probability of there being a complete object in a local window centered on each pixel.¹⁶

- Randomly sample N windows over the image
- Assign each window w a probability score $P(w)$ to indicate its objectness
- Sum all the probability scores in windows that contains pixel x

$$O_p(x) = \sum_{w \in W \text{ and } x \in w} P(W_x) \quad (7)$$

¹⁶Jiang, Peng and Ling, Haibin, “Salient region detection by ufo: uniqueness, focusness and objectness”, in ICCV, 2013



⑥ Smoothness Prior

Saliency
Detection

Background
Development
Features
Priors

Aggregation

Smoothness constraint is often encoded by adding a pair-wise potential to the energy function which encourages neighboring pixels in the image to take the same label.¹⁷

$$w_{ij} = \exp\left(-\frac{\|c_i - c_j\|}{2\sigma_w^2}\right) \quad (8)$$

$$E(S) = \sum_i (S(i) - S_{in}(i))^2 + \lambda \sum_{i,j} w_{ij} (S(i) - S(j))^2 \quad (9)$$

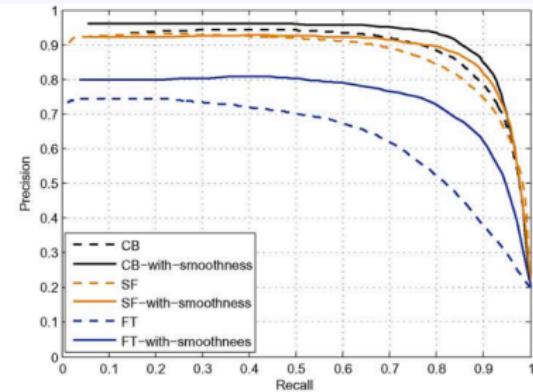
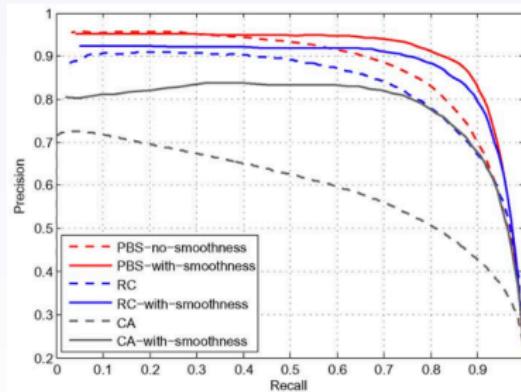
¹⁷Yang, Chuan and Zhang, Lihe and Lu, Huchuan, "Graph-regularized saliency detection with convex-hull-based center prior", in Signal Processing Letters, IEEE, 2013.



⑥ Smoothness Prior

Saliency
Detection

Background
Development
Features
Priors
Aggregation

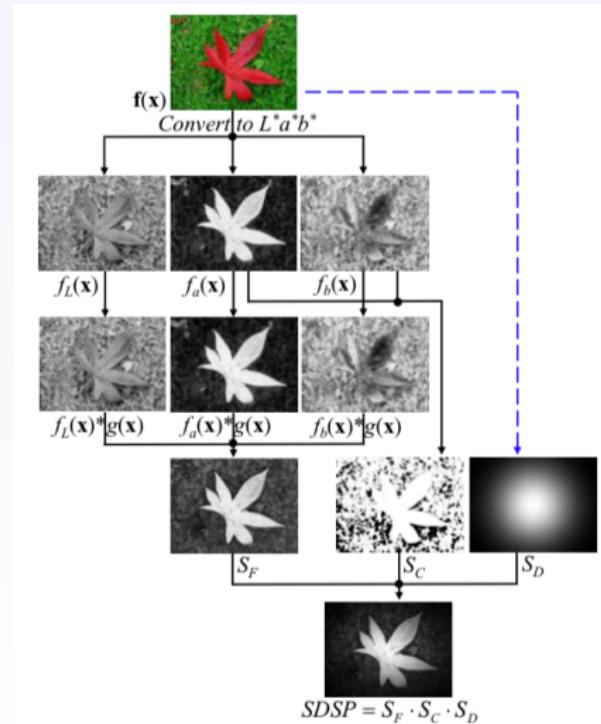




Example: SDSP¹⁸ (**Demo**)

Saliency
Detection

Background
Development
Features
Priors
Aggregation



¹⁸Zhang, Lin and Gu, Zhongyi and Li, Hongyu, "SDSP: A novel saliency detection method by combining simple priors", in ICIP, 2013.



① Non-linear¹⁹

Saliency
Detection

Background
Development
Features
Priors
Aggregation

Combining the focusness, the objectness and the uniqueness maps:

$$S = \exp(F + U) \times O \quad (10)$$

¹⁹Peng Jiang et al., “Salient region detection by ufo: Uniqueness, focusness and objectness”, in ICCV, 2013.



② Energy minimization²⁰

Saliency
Detection

Background
Development
Features

Priors

Aggregation

$$\underbrace{\sum_{i=1}^N w_i^{bg} s_i^2}_{background} + \underbrace{\sum_{i=1}^N w_i^{fg} (s_i - 1)^2}_{foreground} + \underbrace{\sum_{i,j} w_{ij} (s_i - s_j)^2}_{smoothness} \quad (11)$$

²⁰Wangjiang Zhu and Shuang Liang, "Saliency Optimization from Robust Background Detection", in CVPR, 2014.



③ Machine Learning²¹

Saliency
Detection

Background

Development

Features

Priors

Aggregation

Training examples:

- a set of confident regions $R = \{R_1, R_2, \dots, R_Q\}$
- the responding saliency scores $A = \{a_1, a_2, \dots, a_Q\}$

²¹Huaizu Jiang et al., “Salient object detection: a discriminative regional feature integration approach”, in CVPR, 2013.



Saliency
Detection

Background
Development
Features
Priors
Aggregation

Thanks

Yafei Zhu
Ocean University of China
2015.03