ACM International Conference on Internet Multimedia Computing and Service

How Important is Location in Saliency Detection

Tongwei Ren[1], Ran Ju[1], Yan Liu[2], Gangshan Wu[1]

[1] State Key Laboratory for Novel Software Technology, Nanjing University

[2] Department of Computing, The Hong Kong Polytechnic University

July 10, 2014

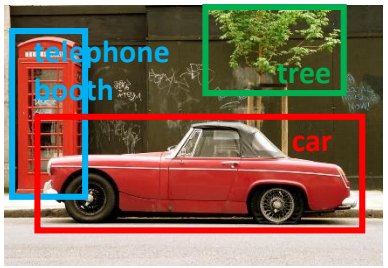
Outline

- **Introduction**
- **Overview**
- **Experiment**
- **Conclusion**

Saliency Detection

- Detect the regions attracting human attention from image content

- Used as a fundamental of many multimedia applications, such as salient object recognition, information retrieval, adaptive compression, and content-aware editing



salient object recognition

information retrieval

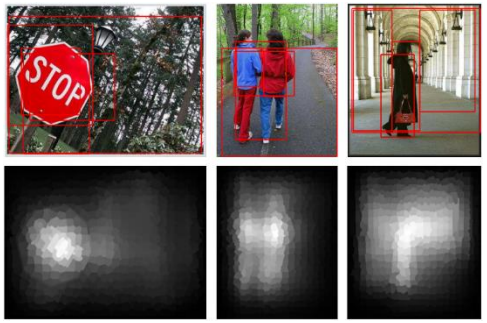
adaptive compression

content-aware editing

Current Methods

- **Low-level features based methods**
  - **Work in a bottom-up manner by integrating low-level visual features**
  - **Mechanism of human eye fixation is still unclear**
- **High-level object based methods**
  - **Detect the interesting objects and assign high saliency**
  - **Limited by the object detection performance**

[Itti, TPAMI'98]



[Jia, ICCV'13]

Location Information

- Only studied in few previous works
  - [Liu, CC'03] applied location in CT image analysis
  - [Judd, ICCV'09] trained the classifier on eye fixation location
  - [Liu, TPAMI'11] considered color spatial distribution
- One reason to avoid using location information is salient objects may appear in any location in some special applications, such as surveillance
- Location provides very useful information in more common applications, for example, detecting salient regions in natural images

Outline

- Introduction
- Overview
- Experiment
- Conclusion

Basic Idea

- **Objective**
  - **Reveal the importance of location information in saliency detection in natural images**

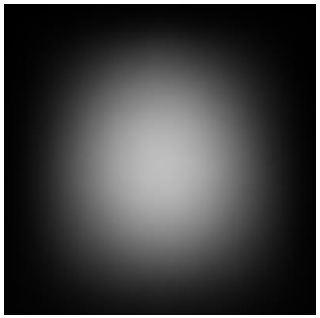
- **Intuition**
  - **Salient objects are usually placed in the center or golden section ratio of image in photography**



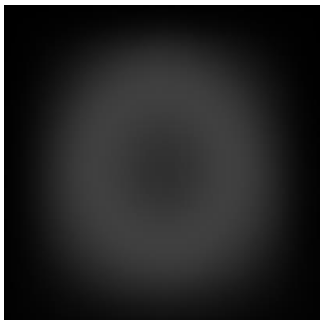

Observation of THUS10000

- **THUS10000 dataset includes 10000 images with pixel-level manually labeled saliency maps**

- **Calculate the mean value and variance of all the saliency maps**



mean value



variance

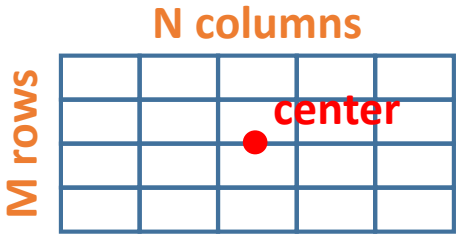
Outline

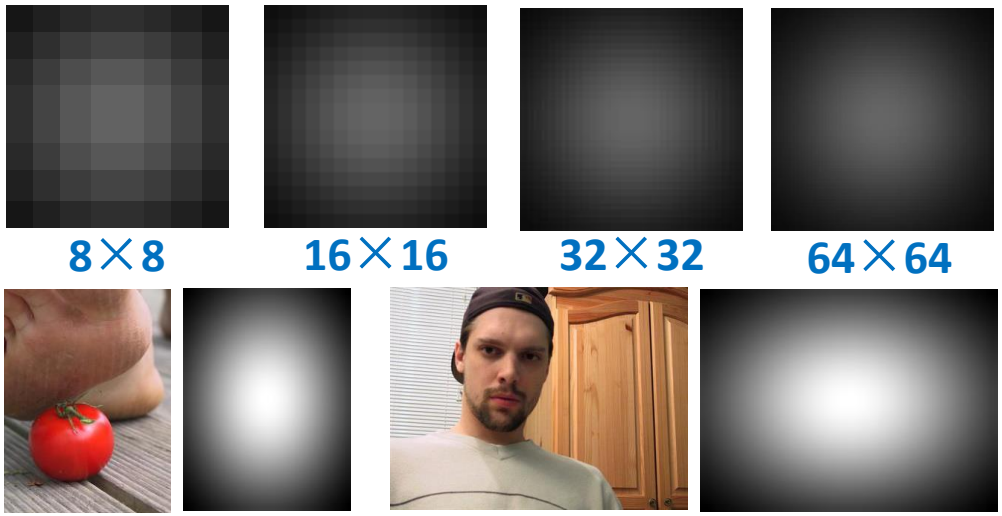
- Introduction
- Overview
- Experiment
- Conclusion

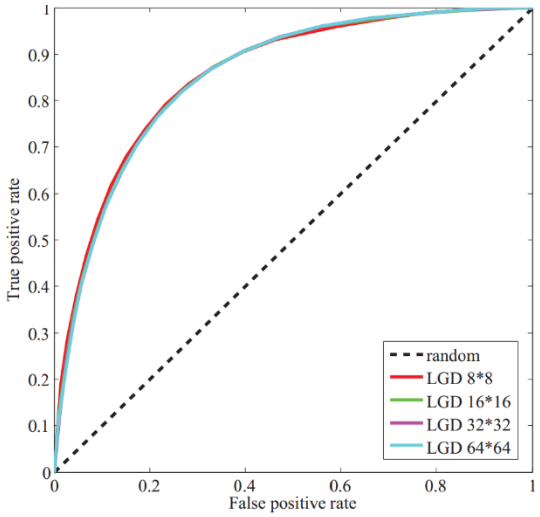
Location based Gaussian Distribution (LGD)

- **Decompose the image into M×N patches, and assign the saliency value to each patch $p_{m,n}$ based on its normalized distance to the center of image**

$$s_{m,n} = \frac{1}{2\pi\sigma^2} e^{-\frac{(m'-1)^2+(n'-1)^2}{2\sigma^2}}$$







Location based Saliency Propagation (LSP)

- **Initialized with LGD and propagate the saliency among the patches**

$$s'_{i,j} = \omega(:, p_{i,j})^T \mathbf{s}$$

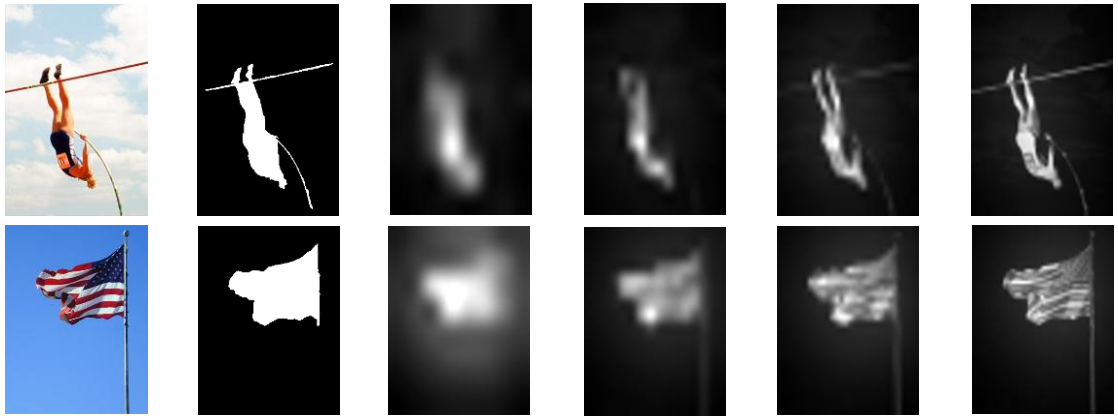
$$\omega(:, p_{i,j}) = [\omega(p_{1,1}, p_{i,j}), \cdots, \omega(p_{M,N}, p_{i,j})]^T$$

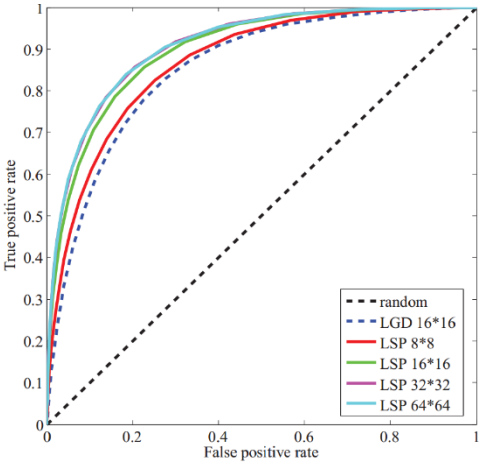
$$\mathbf{s} = [s_{1,1}, \cdots, s_{M,N}]^T$$

$$\omega(p_{m,n}, p_{i,j}) = \omega_s(p_{m,n}, p_{i,j}) \cdot \omega_C(p_{m,n}, p_{i,j})$$

$$\omega_s(p_{m,n}, p_{i,j}) = e^{-\frac{(m'-i')^2+(n'-j')^2}{\sigma^2}}$$

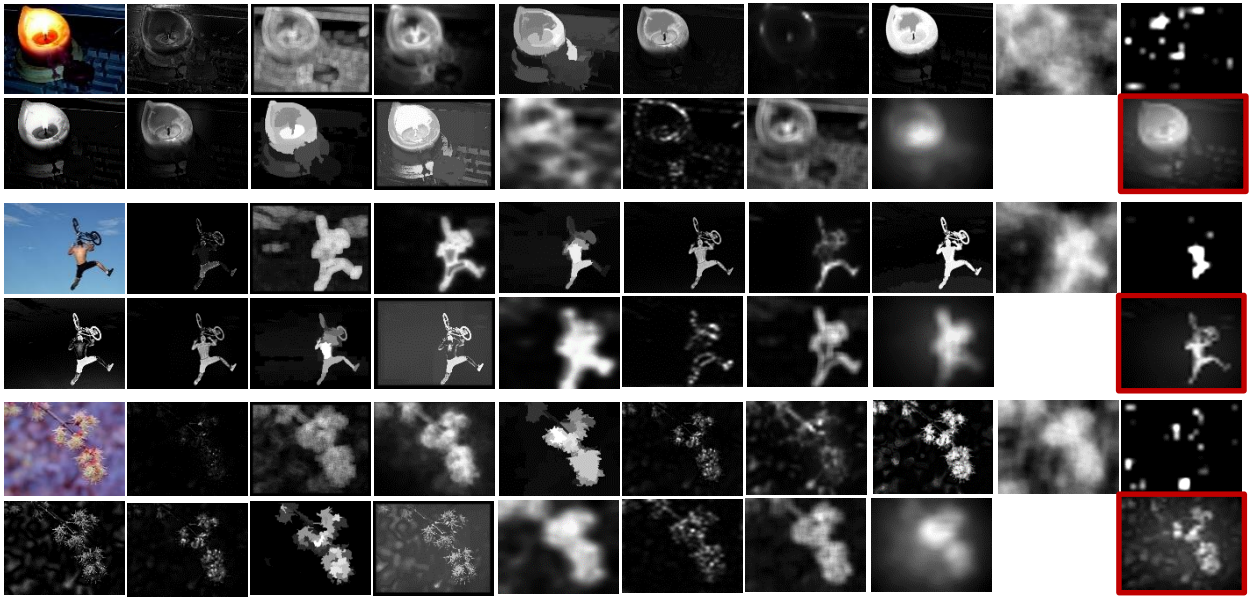
$$\omega_C(p_{m,n}, p_{i,j}) = 1 - \|c_{m,n} - c_{i,j}\|_2$$

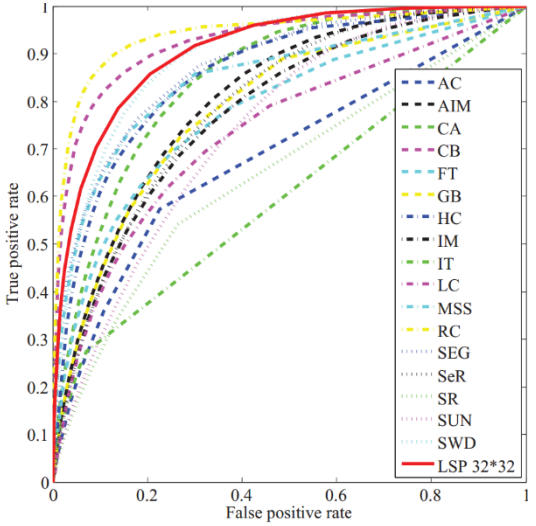




Comparison

- **Compared to 17 saliency detection methods**
  - **AC**[ICVS'08], **AIM**[JV'09], **CA**[CVPR'10], **CB**[BMVC'11], **FT**[CVPR'09], **GB**[NIPS'06], **HC**[CVPR'11], **IM**[CVPR'11], **IT**[TPAMI'98], **LC**[MM'06], **MSS**[ICIP'10], **RC**[CVPR'11], **SEG**[ECCV'10], **SeR**[JV'09], **SR**[CVPR'07], **SUN**[JV,08], **SWD**[CVPR'11]





What is the Limitation?

- **LSP performance will be unstable when image content is complex or object is similar to background**

Good results



Bad results



Degenerate to LGD

Cannot distinguish salient region and background

Confuse object and background

Outline

- Introduction
- Overview
- Experiment
- Conclusion

Contribution

- **Reveal the importance of location information in saliency detection**
- **Validate it with two location based saliency detection approaches, which completely ignore image content or only use weak assistance of image content**
- **Provide another possibility to efficiently and accurately detect saliency in natural images**

Thank You

rentw@nju.edu.cn