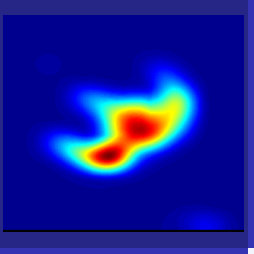


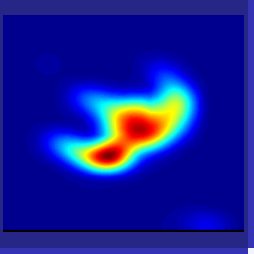
Paper Sharing

Bottom-Up Saliency Detection Model Based on Human Visual Sensitivity and Amplitude Spectrum

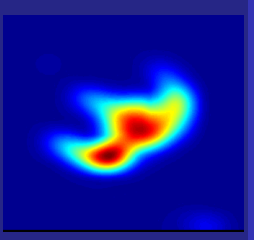
Zhao Hongmiao

CVBIOUC
Ocean University of China
http://vision.ouc.edu.cn/~zhenghaiyong

December 26, 2014

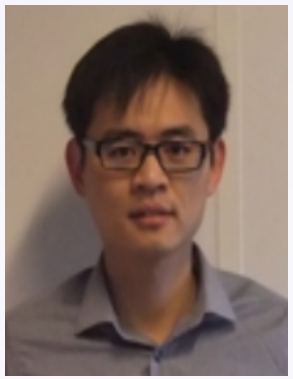


Contents

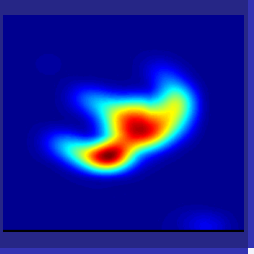


Introduction

- Author
  - Fang Yuming
  - 1984.10.
  - lecturer
  - Jiangxi University of Finance and Economics



- Citation
  - IEEE TRANSACTIONS ON MULTIMEDIA
  - Impact Factor: 1.776
  - Citation Rate: 31
- Highlights
  - The human visual sensitivity and QFT amplitude spectrum
  - Application: image retargeting



Saliency Detection Model

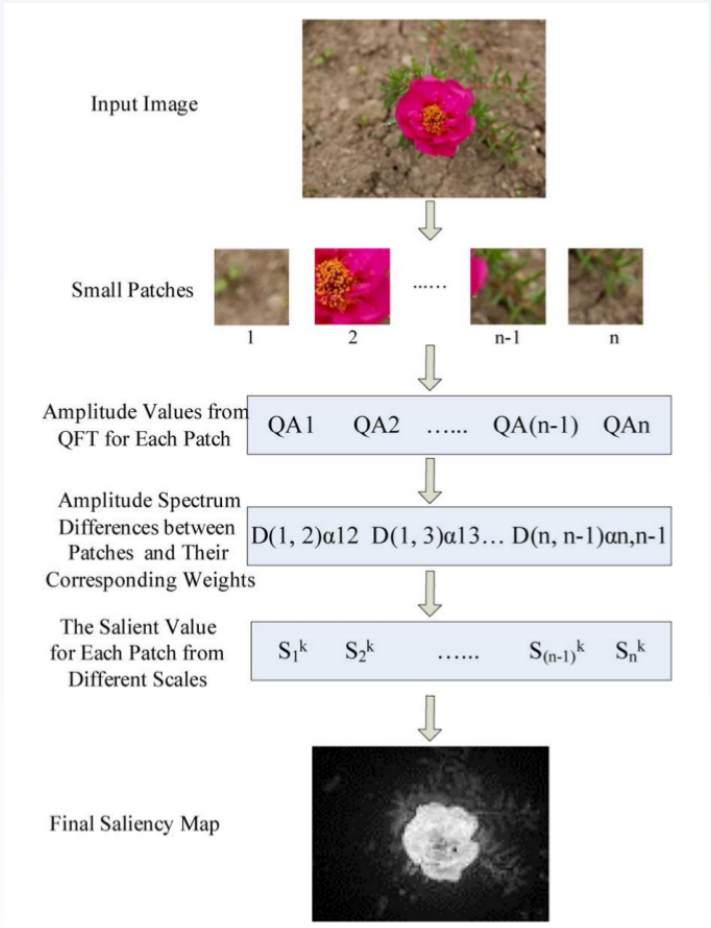
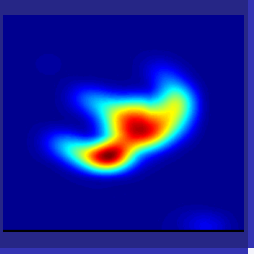
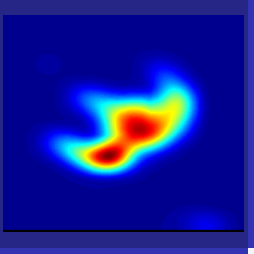


Figure : Framework of the proposed saliency detection model.



Saliency Value for Each Patch

- Two factors
  - The patch differences between this image patch and all other image patches in the input image$\Longrightarrow \mathcal{D}_{(i,j)}$
  - The weighting for these patch differences$\Longrightarrow \alpha_{ij}$
- Definition
  - $S_i = \sum_{j \neq i} \alpha_{ij} \mathcal{D}_{(i,j)}$



Differences Between Image Patches and Their Weighting to Saliency Value

- Differences Between Image Patches
  - Quaternion Fourier Transform
  - $\mathcal{D}_{(i,j)} = \sqrt{\sum_m (log(\mathcal{A}_m^i + 1) - log(\mathcal{A}_m^j + 1))^2}$
- The weight for the patch difference between different patches[1]

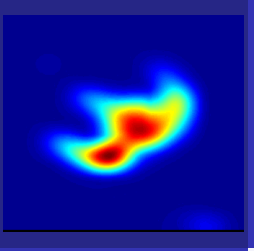
$$\mathcal{C}_s(f, e) = \frac{1}{\mathcal{C}_t(f, e)} \tag{1}$$

$$\mathcal{C}_t(f, e) = \mathcal{C}_0 exp(\alpha f \frac{e + e_2}{e_2}) \tag{2}$$

$$e = tan^{-1}(\frac{d}{v}) \tag{3}$$

$$\alpha_{ij} = \frac{1}{\mathcal{C}_0 exp(\alpha f \frac{e+e_2}{e_2})} \tag{4}$$

[1] W.S. Geisler and J.S. Perry, "A real-time foveated multi-solution system for low-bandwidth video communication," in Proc. SPIE, Jul. 1998, vol. 3299, pp. 294–305.



Patch Size and Scale for Final Saliency Value

- Patch Size

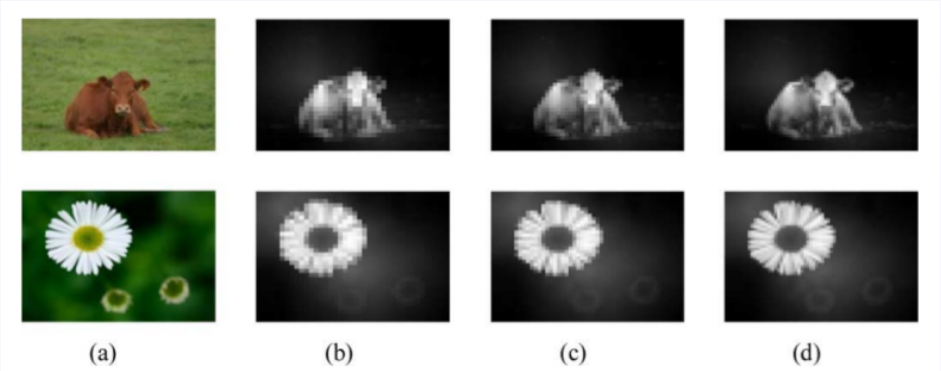
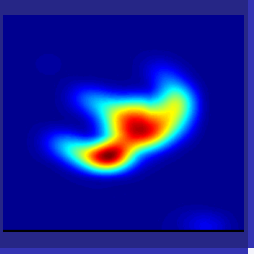


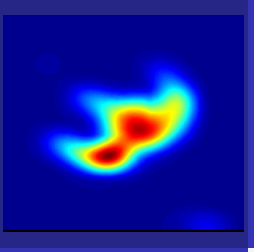
Figure : Original images and its different saliency maps with different patch sizes.

- The smaller image patch size, the more accurate saliency map we can get.
- Suitable patch size: consideration of fovea characteristics, saliency detection performance, and the computational complexity.



Patch Size and Scale for Final Saliency Value

- Scale
  - the steerable pyramid algorithm
  - $S_i = \frac{1}{N} \sum_k S_i^k$



Evaluations

- Database: MSRA[2]
- Comparison methods: PO, SR, MD, FG (temporary abbreviation)

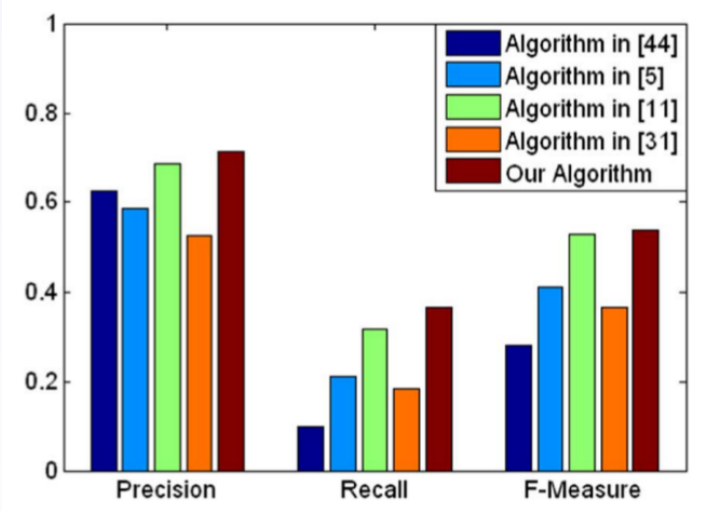
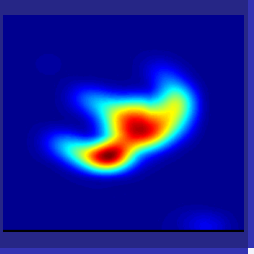


Figure : Experiment results for the comparison between the proposed model and others.

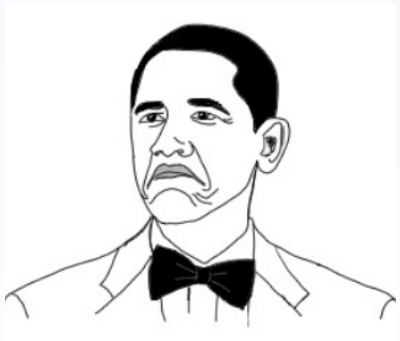
---

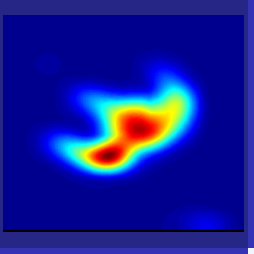
[2]T. Liu, J. Sun, N. Zheng, X. Tang, and H. Y. Shum, "Learning to detect a salient object," in Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, 2007.



Conclusions

- Advantages
  - Image Patch $\Longrightarrow$ Superpixels
  - $\mathcal{D}_{(i,j)}$ is novel
  - Steerable pyramid algorithm
- Disadvantages
  - Accuracy
  - Computational complexity

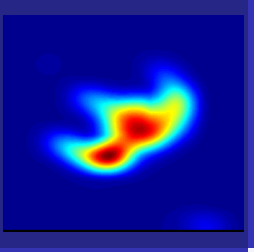




Thinking

- Conclusion in frequency domain

Model	Method	Model	Method
SR	Residual Spectrum	HFT	SSS
PFT	Only Phase Spectrum	SDSP	log-Gabor
SIG	DCT+Sign	FT	DoG




Thank you