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, Nanjing, China
[2] Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China
rentw@nju.edu.cn, juran@smail.nju.edu.cn, csyliu@comp.polyu.edu.hk, gswu@nju.edu.cn

ABSTRACT

Current saliency detection methods mainly work on exploring the potential of low-level and high-level visual features, such as color, texture and face, but treat location information as a weak assistance or completely ignore it. In this paper, we reveal the importance of location information in saliency detection. We analyze the largest public image dataset for saliency detection *THUS10000*, and find the relationship between content location and saliency distribution. To further validate the effect of location information, we propose two location based saliency detection approaches, location based Gaussian distribution and location based saliency propagation, which make use of no or weak assistance of image content. Experimental results show that location based saliency detection can obtain much better performance than random selection, even better than most state-of-the-art saliency detection methods.

Categories and Subject Descriptors

I.2.10 [**Artificial Intelligence**]: Vision and Scene Understanding; I.4.9 [**Image Processing and Computer Vision**]: Applications

General Terms

Algorithms, Human Factors

Keywords

Saliency detection, location information, patch representation, saliency propagation

1. INTRODUCTION

Saliency detection, i.e. detecting the regions attracting human attention from image content, plays an important role in many vision and multimedia tasks [4]. It usually generate the salient regions without high level processing, and supplies a better allocation of computing resource. Thus, saliency detection has been used as a fundamental of many multimedia applications, such as image classification [13] and information retrieval [14].

In the past years, various saliency detection methods have been proposed. The existing methods mainly focus on exploring the potential of some low-level features, such as color and texture, in saliency detection. Itti *et al.* [5] first proposed a computational model for gaze prediction, and Borji *et al.* made an excellent summary [3]. These methods works in a bottom-up manner by integrating early visual features across multiple scales, which aims at predict the fixation point and tends to highlight edges or corners of objects. A few recent works try to detect the entire salient objects based on early visual information like color, texture etc. Achanta *et al.* [1] propose to centering the intensities of whole images to highlight the salient objects. Liu *et al.* [10] made a combination of local and global color features to infer the salient object. Cheng *et al.* [4] employ global color contrast for salient object detection, which shows both high accuracy and efficiency in the *THUS10000* dataset. Margolin *et al.* [11] try to measure the pattern differences of patches using PCA and achieves encouraging results in five open datasets [1, 2, 10, 12].

Besides the low-level features, some saliency detection methods also pay attention to making use of some high-level visual information. In [6], Jia *et al.* propose to combine high level saliency priors by objectness measurement with low level appearance models. Jiang *et al.* [7] explore uniqueness, focusness and objectness for salient region detection. These methods, though have limitations in particular images, achieve satisfactory results in general open datasets which are consisted of natural images.

Compare to the above mentioned low-level and high-level features, location information has been only studied in few previous works [8, 9, 10]. One reason to avoid using location information in saliency detection is salient objects may appear in any location in some special applications, such as surveillance. But in more common applications, for example, detecting salient regions in natural images, location provides useful information.

In this paper, we focus on revealing the importance of location information in saliency detection. We first analyze the relationship between location and saliency distribution on *THUS10000* dataset, which includes 10000 images with pixel-level manually labeled saliency maps. We calculate the mean value and variance of all the saliency maps and analyze the results. Furthermore, we propose two location based saliency detection, including location based Gaussian

ICIMCS '14 Jul. 10-12, Xiamen, Fujian, China

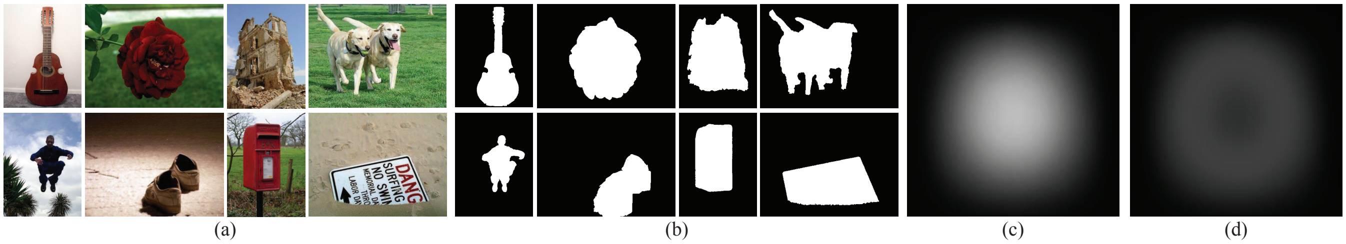


Figure 1: Analysis of *THUS10000* dataset. (a) Example of source images. (b) Corresponding manually labeled saliency maps. (c) Mean of all saliency maps. (d) Variance of all saliency maps.

distribution and location based saliency propagation, and further validate the effect of location information in saliency detection on *THUS10000* dataset.

2. RELATIONSHIP BETWEEN LOCATION AND SALIENCY DISTRIBUTION

Location information plays an important role in salient content representation. Inspired by the characteristics of human visual perception, salient objects are usually placed in the center or golden section ratio of image in photography.

To validate the effect of location information in saliency representation, we make statistics on a public image dataset *THUS10000* [4]. To the best of our knowledge, it is the largest dataset with various source images for saliency detection (Figure 1(a)). Though another saliency detection image dataset *THUS15000* provided by the same authors includes more source images, its source images are limited in five categories. To each source image, $THUS10000$ provides a pixel-level manually labeled saliency map (Figure 1(b)). We resize all the saliency maps to square and calculate the mean value and variance of the resized saliency maps. As shown in Figure 1(c) and (d), we can find that the region near to image center has high mean value and low variance value, which means the centric regions of most images are assigned to high saliency values. On the contrary, the region far from image center has low mean value and variance value, which means these regions are always treated as not salient in detection.

3. LOCATION BASED SALIENCY DETECTION

3.1 Location based Gaussian Distribution

Inspired by Figure 1(c), we utilize content-independent location based Gaussian distribution to detect saliency regions. We decompose the image into $M \times N$ patches, i.e., the number of patches is MN, and assign the saliency value to each patch $p_{m,n}$ based on its normalized distance to the center of image:

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

where $m' = \frac{2m}{M+1}$ and $n' = \frac{2n}{N+1}$ are the normalized coordinate of patch $p_{m,n}$ to the center $(\frac{M+1}{2}, \frac{N+1}{2})$; σ is a parameter to adjust saliency distribution, and $\sigma^2 = 0.4$ in our experiments.

Based on Eq. (1), we obtain the initial saliency map, and further normalize its value to the range of [0, 1] and resize it to source image size.

3.2 Location based Saliency Propagation

Location based Gaussian distribution emphasizes the importance of location information but completely ignores image content, which may limit the performance of saliency detection. In fact, the images with different content should have different saliency maps even the salient regions may locate near to the image centers. Hence, we bring image content information in location based saliency detection.

To keep the dominant role of location information in saliency detection, we only make a weak use of image content, i.e., propagating the saliency between the patches with similar mean color. We initialize the saliency value of each patch $p_{i,j}$ with Eq. (1) and propagate the saliency from all the patches to $p_{i,j}$:

$$s_{i,j}^{'} = \boldsymbol{\omega}(:, p_{i,j})^T \mathbf{s}, \tag{2}$$

where $\boldsymbol{\omega}(:, p_{i,j}) = [\omega(p_{1,1}, p_{i,j}) \dots \omega(p_{M,N}, p_{i,j})]^T$ is the propagation weight vector of all the patches to patch $p_{m,n}$, and $\mathbf{s} = [s_{1,1} \dots s_{M,N}]^T$ is the saliency vector composed of the saliency values of all the patches.

We define the propagation weight based on the normalized spatial distance and color similarity between two patches:

$$\omega(p_{m,n}, p_{i,j}) = \omega_s(p_{m,n}, p_{i,j}) \cdot \omega_c(p_{m,n}, p_{i,j}). \tag{3}$$

Here, $\omega_s(p_{m,n}, p_{i,j})$ is the weight based on spatial distance between patch $p_{m,n}$ and $p_{i,j}$:

$$\omega_s(p_{m,n}, p_{i,j}) = e^{-\frac{(m'-i')^2+(n'-j')^2}{\sigma^2}}, \tag{4}$$

where σ is a parameter to adjust saliency distribution, and (m', n') and (i', j') are the normalized coordinates of patch $p_{m,n}$ and $p_{i,j}$ same to Eq. (1).

And $\omega_c(p_{m,n}, p_{i,j})$ is the weight based on the distance between the mean color values of patch $p_{m,n}$ and $p_{i,j}$:

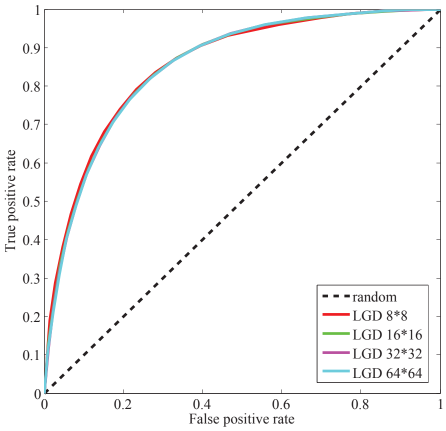
$$\omega_c(p_{m,n}, p_{i,j}) = 1 - ||\mathbf{c}_{m,n} - \mathbf{c}_{i,j}||_2, \tag{5}$$

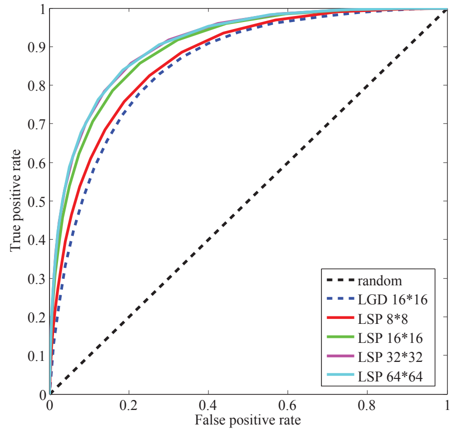
where $\mathbf{c}_{m,n}$ and $\mathbf{c}_{i,j}$ are the mean color values of patch $p_{m,n}$ and $p_{i,j}$ in $L^*a^*b^*$ color space.

To make the saliency propagation convergence, we normalize the propagation weights of each patch to other patches, and conserve the saliency in propagation:

$$\omega^*(p_{m,n}, p_{i,j}) = \frac{\omega(p_{m,n}, p_{i,j})}{\sum_{i=1}^{M} \sum_{j=1}^{N} \omega(p_{m,n}, p_{i,j})}. \tag{6}$$

Based on Eq. (2)-(6), saliency map is iteratively updated by propagating saliency among patches till the change of saliency map is less than a pre-defined threshold, which equals $\frac{1}{MN}$ in our experiments.





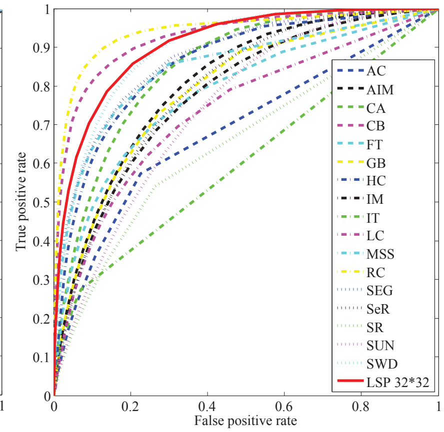


Figure 2: ROC curve for LGD using different patch decomposition.

Figure 3: ROC curve for LSP using different patch decomposition.

Figure 4: Comparison of ROC curves for different saliency detection methods.

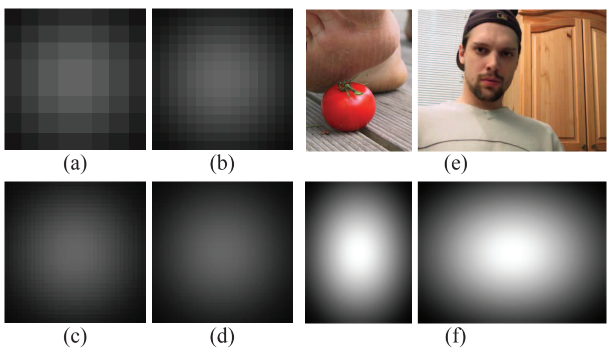


Figure 5: Example of saliency maps generated by LGD using different patch decomposition. (a)-(d) Initial saliency maps using different patch decomposition in 8×8, 16×16, 32×32 and 64×64. (e) Source images. (f) Saliency maps generated by LGD using patch decomposition in 32×32.

4. EXPERIMENTS

4.1 Dataset and Experiment Setting

We validate the location based saliency detection approaches on *THUS10000* dataset, including location based Gaussian distribution (LGD) and location based saliency propagation (LSP). To each approach, we evaluate its performance under four kinds of patch settings, including the pixel numbers of 8×8, 16×16, 32×32 and 64×64. All the experiments were implemented in Matlab on a computer with 3.4GHz CPU and 8GB memory.

4.2 Experimental Results

Figure 5 shows the examples of saliency maps generated by LGD approach. In LGD approach, the initial saliency map is generated by based on the normalized distance to image center, and normalized to the value range [0, 1] and resized to image size. So the saliency maps generated using different patch decomposition are similar. Figure 2 and Table 1 show the performance of LGD using different patch decomposition. It shows that different patch decomposition has little influence to saliency detection performance. And it is obvious that location information has positive effect to saliency detection even completely ignoring image content.

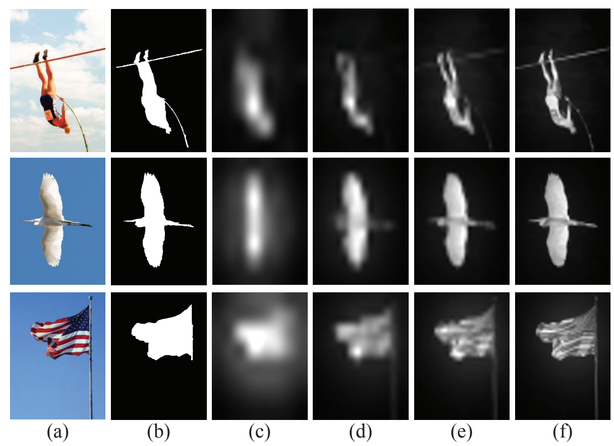


Figure 6: Example of saliency maps generated by LSP using different patch decomposition. (a) Source image. (b) Manually labeled groundtruth. (c)-(f) Saliency maps generated by LSP using patch decomposition in 8×8, 16×16, 32×32 and 64×64.

Figure 6 shows the examples of saliency maps generated by LSP approach. Different to LGD approach, the number of patches influences the saliency detection performance in LSP approach. The increasing of patch number makes the patch size smaller, which benefits to obtain more accurate salient region boundaries, such as the top and middle rows in Figure 6. Meanwhile, the color values within each patch will be more similar, which may lead to the mean color values of patches are more distinct and prevent the saliency propagation among patches, such as the bottom row in Figure 6. Figure 3 and Table 1 show the performance of LSP approach using different patch decomposition. We can find that the performance of saliency detection is improved when increasing patch number from 8×8 to 32×32, but it doesn't change when further increasing patch number.

We also compared the location based saliency detec-

Table 1: **AUC of ROC curves for LGD and LSP using different patch decomposition.**

	LGD	LSP
8×8	0.85	0.87
16×16	0.85	0.90
32×32	0.85	0.91
64×64	0.85	0.91

tion approaches with the state-of-the-art saliency detection methods. All the saliency detection results of 17 compared methods are provided by *THUS10000* dataset [4]. Different to only location information is used in our approach, all the features are allowed to use in these saliency detection methods. Figure 4 shows the comparison result of RoC curves for different saliency detection methods. It shows that the performance of LSP approach is better than some state-of-the-art saliency detection methods.

4.3 Discussion

In the experiments, we also find some limitation of location based saliency detection. When the salient region and background both have complex structures, the performance of LSP approach will degenerate to LGD, such as the top row in Figure 7, especially when the number of patch is large. The reason is only mean value of each patch may be distinct to each other and the saliency propagation from each patch to others become uniform. Another situation of our drawback is shown in the bottom row of Figure 7. For only very simple image content feature, mean color value of each patch, is used to determine the propagation weight, the salient region cannot be distinguished from the background. The saliency of center region is propagated to the similar background part, and the rest part of background is detected as salient regions.

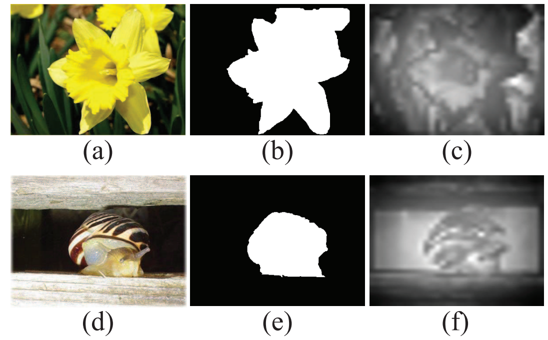


Figure 7: Examples of our drawback. (a) and (d) Source images. (b) and (e) Manually labeled saliency maps. (c) and (f) Saliency maps generated by LSP using patch decomposition in 32×32.

5. CONCLUSIONS

In this paper, we reveal the importance of location information in saliency detection. Based on the analysis of relationship between image content location and saliency distribution, we find that location information has obvious influence to saliency. Furthermore, we propose two location based saliency detection approach, which completely ignores image content and uses weak assistance of image content, respectively. It shows that the location based saliency detection methods can obtain much better performance than random selection, even better than some state-of-the-art saliency detection methods.

6. ACKNOWLEDGMENTS

This paper is supported by the National Science Foundation of China (No. 61321491, 61202320), Research Project of Excellent State Key Laboratory (No.61223003), Natural Science Foundation of Jiangsu Province (No.BK2012304), and National Special Fund (No.2011ZX05035-004-004HZ).

7. REFERENCES

[1] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-tuned salient region detection. In *CVPR*, pages 1597–1604. IEEE, 2009.

[2] S. Alpert, M. Galun, R. Basri, and A. Brandt. Image segmentation by probabilistic bottom-up aggregation and cue integration. In *CVPR*, pages 1–8. IEEE, 2007.

[3] A. Borji, H. R. Tavakoli, D. N. Sihite, and L. Itti. Analysis of scores, datasets, and models in visual saliency prediction. In *ICCV*, pages 921–928. IEEE, 2013.

[4] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu. Global contrast based salient region detection. In *CVPR*, pages 409–416. IEEE, 2011.

[5] L. Itti, C. Koch, E. Niebur, et al. A model of saliency-based visual attention for rapid scene analysis. *IEEE TPAMI*, 20(11):1254–1259, 1998.

[6] Y. Jia and M. Han. Category-independent object-level saliency detection. In *ICCV*, pages 1761–1768. IEEE, 2013.

[7] P. Jiang, H. Ling, J. Yu, and J. Peng. Salient region detection by ufo: Uniqueness, focusness and objectness. In *ICCV*, pages 1976–1983. IEEE, 2013.

[8] T. Judd, K. Ehinger, F. Durand, and A. Torralba. Learning to predict where humans look. In *ICCV*, pages 2106–2113. IEEE, 2009.

[9] T. Liu, S. D. Slotnick, J. T. Serences, and S. Yantis. Cortical mechanisms of feature-based attentional control. *Cerebral Cortex*, 13(12):1334–1343, 2003.

[10] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H.-Y. Shum. Learning to detect a salient object. *IEEE TPAMI*, 33(2):353–367, 2011.

[11] R. Margolin, A. Tal, and L. Zelnik-Manor. What makes a patch distinct? In *CVPR*, pages 1139–1146. IEEE, 2013.

[12] V. Movahedi and J. H. Elder. Design and perceptual validation of performance measures for salient object segmentation. In *CVPR Workshops*, pages 49–56. IEEE, 2010.

[13] X. Yang, T. Zhang, and C. Xu. Locality discriminative coding for image classification. In *ICIMCS*, pages 52–55. ACM, 2013.

[14] S. Zhang, Q. Tian, Q. Huang, W. Gao, and Y. Rui. Multi-order visual phrase for scalable image search. In *ICIMCS*, pages 145–149. ACM, 2013.