



A Study on the Salient Object Detection Using the Edge Background and Average Filter

Hong-Ki Min*

This work was supported by Incheon National University Research Grant in 2017.

Abstract

In this paper, we propose salient object detection method based on the edge background and average filter for salient object detection that can be applied to computer vision community. We assume edge pixels are background in input image and photographers tend to frame the salient object near the center of the image. We obtain a saliency value through the method of comparing color information of background pixels which derived from each edge with whole pixel in image. It generates background prior feature maps based on obtained saliency. And we generate saliency map through weighted near the center of background prior feature maps. We compare precision and recall of the proposed method with previous 5th salient object detection methods to test usefulness. The experimental results show that proposed method improves in performance than compared previous saliency object detection methods. On average, the F-Measure and AUC were increased by approximately 7.0%, 3.4%.

요 약

본 논문에서는 컴퓨터 비전 분야에서 활용할 수 있는 돌출 객체 검출을 위하여 배경 경계와 평균 필터 기반의 돌출 객체 검출 방법을 제안한다. 배경은 주어진 이미지의 테두리 픽셀이고 사진작가들이 미학을 이유로 돌출 객체를 주로 이미지의 중앙에 위치시킨다고 가정하였다. 주어진 이미지의 전체 픽셀과 각 경계면으로부터 추출된 배경 픽셀의 색상 정보를 비교하는 방법을 통해 각 픽셀의 돌출 값을 구한다. 구해진 돌출 값을 기준으로 배경 우선 특징맵을 생성한다. 그리고 배경 우선 특징맵의 중심 부분에 가중치를 부여하여 돌출맵을 생성한다. 제안하는 방법의 유용성을 확인하기 위하여 기존 5개의 검출 방법들과 정확도 및 재현율을 비교하였다. 실험 결과, 제안하는 방법이 이전의 돌출 객체 검출 방법에 비해 성능이 향상되었음을 확인하였다. 평균적으로 F-Measure는 7.0%, AUC는 3.4%가 증가하였다.

Keywords

salient object, salient map, feature map, computer vision

* Dept. of Information and Telecommunication Eng.
Incheon National University
- ORCID: <http://orcid.org/0000-0001-5043-2432>

· Received: Nov. 20, 2017, Revised: Dec. 16, 2017, Accepted: Dec. 19, 2017
· Corresponding Author: Hong-Ki Min
Department of Information and Telecommunication Eng.
Incheon National University, Korea,
Tel.: +82-32-835-8284, Email: hkmin@inu.ac.kr

I. Introduction

Human beings specifically and selectively recognize and process prominent visual stimuli that feature differences in color, movement direction, and position by focusing on a given image. This capability have been studied by cognitive psychology, neuroscience, and computer vision community mainly because it helps find the objects or regions that efficiently represent a scene and utilize complex vision problems such as scene understanding. We called the salient object which is visually focused by humans [1]. Generally, human focused at a salient object when they get the image. Representative study is saliency map that contains information of the salient regions proposed by Itti et al [2]. Since then, many studies try to predict with saliency map to detect salient object and understand visual attention of human. Bottom-up saliency uses intrinsic cues of image such as brightness, color, contrast, and texture. However, top-down saliency use extrinsic cues of image which made by learn about relationship of images similar [1]-[3].

However, conventional methods for detecting salient objects have many problems. The typical problems are as follows. First, a disadvantage is that excessive saliency appears near the boundary, and the color value in the salient region is attenuated. Second, some methods have the advantage of good boundary detection; however, a problem with this method is that it not only detects the boundary of the salient object but also detects many boundaries of the background area. Third, in some methods, the background area is effectively removed, but the saliency value of the salient region is lowered due to the excessive removal of the background. To solve such problems, this paper proposes a salient object detection method based on the edge background and average filter. In this study, a saliency map was generated using the ASD data set[4]. In order to verify the usefulness of the proposed method, the proposed method was compared

with the saliency maps generated by other salient object detection methods. And, it is confirmed that the proposed method is more useful than compared salient object detection methods.

II. Related Research

Itti et al. explained the human visual search strategies as a feature integration theory. The model used here builds on a second biologically plausible architecture, proposed by Koch and Ullman and at the basis of several models. It is related to the so-called “feature integration theory,” explaining human visual search strategies. The model’s saliency map is endowed with internal dynamics which generate attentional shifts. This model consequently represents a complete account of bottom-up saliency and does not require any top-down guidance to shift attention. The model proposed by Itti uses the Gaussian pyramid for the input image to obtain an image of 9-scales, and the features such as color, brightness, and direction are obtained from the image of each scale obtained. And a feature maps are generated using the difference between center and surround regions and the contrast for each feature. Next, Itti et al. combine the feature maps of each scale, create normalized conspicuity maps, and generated a saliency map by linear combination [2]. Achanta et al. used a frequency-tuned (FT) approach to calculate the saliency map by measuring the difference between the average color value of image and pixel color value of smoothed image [4]. This method detects the boundary of the salient object well, but it also has a limit to detect the boundary of the background area. Cheng et al. proposed a global histogram based contrast method (GC) using a smoothing procedure to reduce color information such as quantized noise for saliency calculations [5]. This method has the disadvantage that only the salient objects are detected well in the case of an image

having a simple background, but the background area is also detected when the background is complicated. Jiang et al. proposed a multi-scale region based method to combine and divide region projection values and obtain a projection map in pixel units [6]. This method is based on three features of the salient object. The three features are that the first salient object is always different from the overall background, the second the salient object is generally located near the center of the image, and the third salient object has well organized boundaries. A saliency map based on these features was calculated. Method of Jiang et al. has a problem that the background region is also detected, although the salient object region exhibits a higher saliency value than the background. Cheng et al. proposed a region based contrast (RC) method that measures the global contrast between all regions of an image. The basic idea is that humans are more interested in the area with high contrast than the surrounds. Also, they thought important for spatial relations. The spatial relation is that the contrast appearing around the region is more saliency than the contrast of the distant location [7].

III. Proposed Method

In this paper, we propose salient object detection method using the edge background[8] end average filter. Proposed method has saliency value of pixel wise, and it uses bottom-up method using intrinsic cues of image. We assume edge pixels are background in input image and used to solve the problem of attenuation of color values in salient objects. Therefore we use background prior and average filter.

Proposed method consists of preprocessing process, background prior feature map generation, average filtered map generation and saliency map generation. In preprocessing process, we cluster all pixels in

input image by EDISON [3]. In background prior feature map generation process, we compute saliency through the method of comparing color information of background pixels obtained from each edge with all pixels in input image. We generate background prior feature maps : F_{ld} , F_{lr} , F_{lt} , F_{lb} of each side reference by obtained saliency of each pixel. Each background prior feature maps generate by the four sides. Each background prior feature maps combine to background feature map F_{bg} by multiplication. In average filtered map generation process, we generate average filtered map by applying the average filter A to the background feature map. And we integrate weighting average filtered map and background prior feature map as the saliency map. This process was shown in Fig. 1.

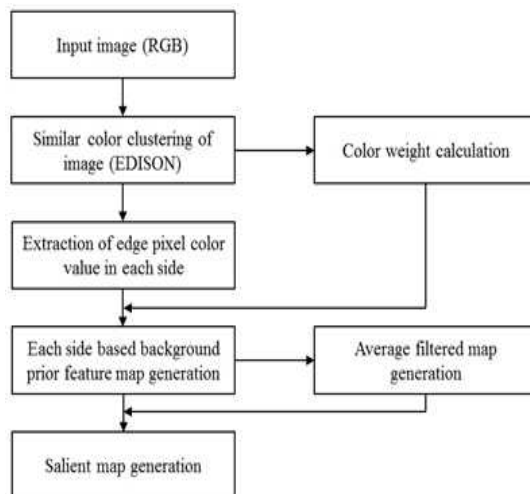


Fig. 1. Flowchart of proposed method

Eq. 1 is a formula of the left background prior feature map F_{ld} . Eq. 2 is a formula of the right background prior feature map F_{lr} . Also top background prior feature map F_{lt} . and bottom background prior feature map F_{lb} . use same way as Eq. 1 or Eq. 2. Eq. 3 is a formula of color weight w_c for background pixels.

$$F_{bl}(i, j) = \begin{cases} I_{ij} \times \left(\frac{i}{n}\right)^2 & \text{if } B_l(k)(1-w_c) \leq C_{ij} \leq B_l(k)(1+w_c), k=1, \dots, kl \\ 0 & \text{if } I_{ij} = B_l \end{cases} \quad (1)$$

$$F_{br}(i, j) = \begin{cases} I_{ij} \times \left(1 - \left(\frac{i}{n}\right)^2\right) & \text{if } B_r(k)(1-w_c) \leq C_{ij} \leq B_r(k)(1+w_c), k=1, \dots, kr \\ 0 & \text{if } I_{ij} = B_r \end{cases} \quad (2)$$

$$w_c = \begin{cases} \frac{|C_{avg} - C_{avgmax}|}{|C_{avg} - C_{avgmax}| + |C_{avg} - C_{avgmin}|} & \text{if } |C_{avg} - C_{avgmax}| - |C_{avg} - C_{avgmin}| \leq 0 \\ \frac{|C_{avg} - C_{avgmin}|}{|C_{avg} - C_{avgmax}| + |C_{avg} - C_{avgmin}|} & \text{otherwise} \end{cases} \quad (3)$$

Where I_{ij} is color value of position (i, j) in image, B_l is pixels of left background, B_r is pixels of right background, n is width of image and k is elements of B_l or B_r . C_{avg} is color average for the image, C_{avgmax} is color average for each row maximum color value in the image and C_{avgmin} is color average for each row minimum color value in the image. Eq. 4 is a formula of the average filter. We used 9×9 as an average mask m . It was better than others.

$$AF(i, j) = \sum_{x=-(m-1)/2}^{(m-1)/2} \sum_{y=-(m-1)/2}^{(m-1)/2} \frac{1}{m^2} I_{ij} \quad (4)$$

$$S(i, j) = F_{bg}(i, j) + AF(i, j) \quad (5)$$

Eq. 5 is a formula of the saliency value. And $F_{bg}(i, j) = F_{bl}(i, j) + F_{br}(i, j) + F_{bt}(i, j) + F_{bb}(i, j)$

IV. Experiments and Results

In this study, a saliency map was generated using the ASD data set. We utilize the Precision and Recall (PR) curve and F-Measure (F_β) and Area Under ROC Curve (AUC) to evaluate the proposed method and other 5th previous methods : GC[5], CB[6], RC[7], SEG[9], FT[4]. In PR curve, we set

the threshold from 0 to 255 for the saliency maps. The F-Measure is the overall performance measurement computed by the weighted harmonic of precision and recall. Also, we compute the AUC which could show the differences among all the methods clearly. Eq. 6 is a formula of the Precision and Recall by a binary mask M with ground-truth G. We are computed average precision and recall values. Eq. 7 is a formula of the F-Measure and β^2 is set to 0.3 to increase the importance of the Precision value [5].

$$Precision = \frac{|M \cap G|}{M}, Recall = \frac{|M \cap G|}{G} \quad (6)$$

$$F_\beta = \frac{(1 + \beta^2) Precision \times Recall}{\beta^2 Precision + Recall} \quad (7)$$

Fig. 2 shows PR curve for proposed method and other methods. The experimental results show that proposed method improves in performance than compared previous saliency object detection methods. On average, the F-Measure and AUC were increased by approximately 7.0%, 3.4%. Therefore, the proposed method were verified to be useful than the conventional salient object detection methods that were compared in the study.

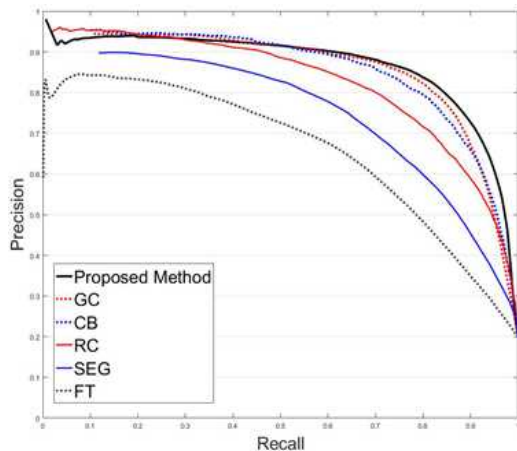


Fig. 2. PR curve of proposed method and other methods

Fig. 3 shows the F-Measure and AUC values for proposed method and other methods. Fig. 4 shows saliency map of proposed method and other methods.

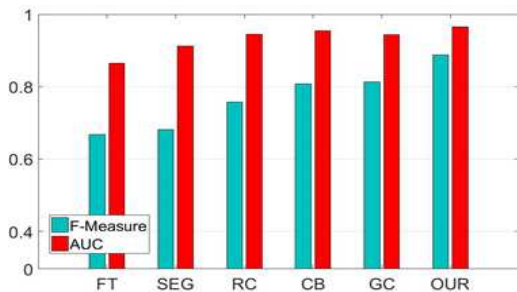


Fig. 3. F-Measure and AUC values of methods

V. Conclusion

In this paper, we proposed salient object detection method using the edge background and average filter. We computed saliency values using proposed method, and generated saliency map. We compare proposed method with other 5th salient object detection methods. The experimental result shows that proposed method improves the performance, demonstrating both higher F-Measure and AUC than compared other methods. However, the background can not be completely removed by comparing the protrusions in pixel units. Therefore, it is necessary to study the method of simultaneously using blocks and pixels for background removal.

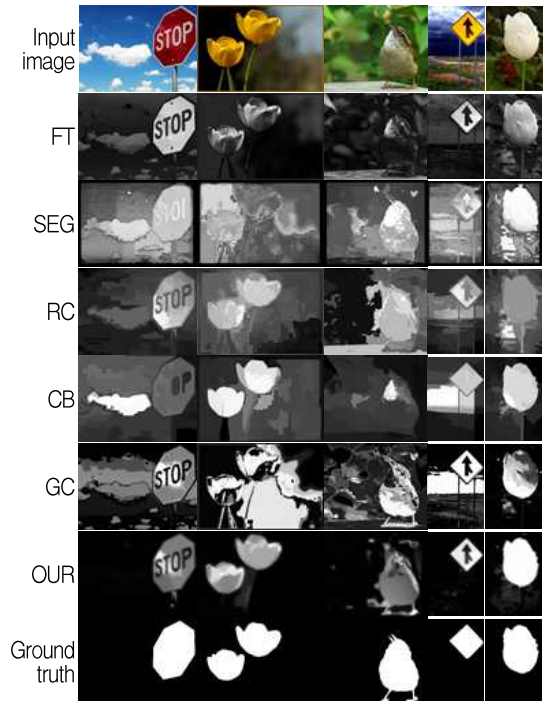


Fig. 4. Saliency map results of methods

References

- [1] A. Borji, M. M. Cheng, H. Jiang, and J. Li, "Salient object detection : A survey", <http://arxiv.org/abs/1411.5878>, Nov. 2014
- [2] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis", *IEEE Trans. Pattern Anal. Intell.*, Vol. 20, No. 11 pp. 1254-1259, Mach. 1998
- [3] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 24, No. 5, pp. 603-619, May 2002
- [4] R. Achanta, S. Hemami, F. Estrada, and S. Süsstrunk, "Frequencytuned salient region detection", in *Proc. IEEE Conf. CVPR*, pp. 1597-1604, June 2009
- [5] M. M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr, and S. M. Hu, "Global contrast based salient region detection", *IEEE Trans. Pattern Anal. Intell.*, Vol. 37, No. 3, pp. 569-582, March 2015.

- [6] H. Jiang, J. Wang, Z. Yuan, T. Liu, and N. Zheng, "Automatic salient object segmentation based on context and shape prior", in *BMVC*, Vol. 6, No. 7, article 9, Sep. 2011.
- [7] M. M. Cheng, G. X. Zhang, N. J. Mitra, X. Huang, and S. M. Hu, "Global contrast based salient region detection", In *CVPR*, pp. 409-416, June 2011.
- [8] J. J. Ryu and H. K. Min, "Saliency Map Generation Based on the Meanshift Clustering and the Pseudo Background", *Journal of Korean Institute of Information Technology*, Vol. 14, No. 11, pp. 119-125, Nov. 2016.
- [9] E. Rahtu, J. Kannala, M. Salo, and J. Heikkilä, "Segmenting salient objects from images and videos", in *Proc. 11th ECCV*, pp. 366-379, Sep. 2010.

Authors

Hong-Ki Min



1979 : BS degree in Dept. of
Electronic Eng., Inha University
1981 : MS degree in Dept. of
Electronic Eng., Inha University
1985 : PhD degree in Dept. of
Electronic Eng., Inha University
1985 ~ 1991 : KIST Senior

Researcher

1991 ~ present : Professor, Dept. of Information and
Telecommunication Eng., Incheon National University

Research interests : HCI, Pattern Recognition, Signal
Process, Rehabilitation Eng.