

Design of Efficient Saliency Map Detector Using Color and Intensity LUT

Changpyo Hong¹, Jong Bin Park¹, Sook Yoon² and Dong Sun Park^{3,4}

¹Department of Electronics Engineering, Chonbuk National University,
Jeonbuk, Korea

²Department of Multimedia Engineering, Mokpo National University,
Jenam, Korea

³Division of Electronics Engineering, Chonbuk National University, Jeonbuk,
Korea

⁴IT Convergence Research Center, Chonbuk National University, Jeonbuk, Korea
leon5678@naver.com, nicedugy@gmail.com, syoon@mokpo.ac.kr,
dspark@jbnu.ac.kr

Abstract

In this paper, we proposed an algorithm which efficiently detects salient objects. First of all, yellow and intensity images are generated from input image and the feature map is generated using a contrast difference between red-green and blue-yellow channel. Secondly, Lookup-Table(LUT) is obtained by histogram of the feature map. Finally, the saliency map is obtained from LUT. The proposed saliency detector shows that this method is able to efficiently extract saliency maps by improving the AUC 10-40%, comparing to existing color-based method and frequency space transform method.

Keywords: Saliency map, Region of interest, Gamma transformation

1. Introduction

In finding a region of interest to understand the image in the field of computer vision, human visual attention has been recognized as a very important issue. Saliency maps from artificial visual attention models are frequently used to effectively find the area of interest at a glance. A saliency map can be obtained from an artificial computer model which mimics biological visual attention of humans. It represents regions in an image that human visual system pay attention to a scene [1].

Detection algorithms for Saliency maps are typically classified into three categories: Bottom-up, Top-down, and Hybrid. A bottom-up type algorithm finds a saliency map without introducing any a priori information but analyzing low-level features from intensity, color, and texture information [2-3]. A top-down type algorithm finds a saliency map with intention to detect a certain set of objects, for example, detecting pedestrian or human face, using specific and higher level features for objects such as shape, color and brightness information [4]. A hybrid type uses a combination of bottom-up and top-down algorithms [5]. The top-down and hybrid methods have shown significant object detection results but these algorithms are very difficult to design and generally very limited to specific applications. Most visual attention researches are currently focused on bottom-up models since they are generally applicable for various purposes.

Itti *et al.* (IT) proposed a renowned bottom-up method which correctly locates salient objects but boundaries of detected objects are not clearly described in the resulting saliency maps. In another work by Xiaodi *et al.* (SR) spectral residues are

used to describe boundaries of objects but positions of objects are not accurately estimated.

In this paper, we propose a bottom-up algorithm that efficiently detects salient objects with clear boundaries and higher accuracy in estimating locations of salient objects. At first, the yellow component and intensity channels are generated from an input color image. The feature map and its corresponding histogram are then produced using a contrast difference between red-green and blue-yellow channel. The histograms of feature map are implemented as Lookup-Table (LUT), and the saliency map is generated by combining elements in the modified LUT. We introduced the opponent color theory and histogram modification and normalization to efficiently extract saliency map by alleviating the disadvantages mentioned above. The proposed detector more likely locates positions of salient objects with clearer boundaries.

This paper is organized as follows: The related research is summarized in Section 2. The proposed saliency detector of algorithm is described in Section 3. Experimental method and results in Section 4. Finally, conclusion is summarized in Section 5.

2. Related Works

Since we propose a bottom-up salient detector, two state-of-the-art bottom-up algorithms are briefly introduced in this section: IT and SR methods.

2.1 IT Method

The IT method separates input image into multiple channels of RGB color, intensity and orientation information. At the second step, it constructs a gaussian pyramid for each channel and performs operation of center surround difference to obtain multiple feature maps from each channel. These feature maps are then combined to generate three conspicuity maps using across-scale combinations and normalization process. Each conspicuity map contains strongly attentive evidence based on color, intensity, and orientation information. The final salient map is generated from the conspicuity maps using pixel-by-pixel winner-take-all operations.

2.2 Spectral Residual (SR) Method

SR method is based on the image statistics which is called as 1/f law, The law states that the average Fourier spectrum of the ensemble of natural images has a distribution of inversely proportional to the frequency. In that sense, a salient object can be viewed as the object which causes abnormal changes in the input frequency spectrum from the average spectrum.

The SR algorithm starts with performing the two-dimensional fast fourier transform (FFT) of the input image. Then the approximated average Fourier spectrum is generated by applying a gaussian smoothing filter to the input spectrum. Spectral residual for salient objects is then calculated by subtracting the modified spectrum from its corresponding input spectrum. The final salient map then be obtained by applying the inverse fast fourier transform.

3. Proposed Algorithm

The proposed algorithm is a bottom-up type based on the opponent colors and intensity of pixels. The overall block diagram of the proposed algorithm is depicted in Figure. 1.

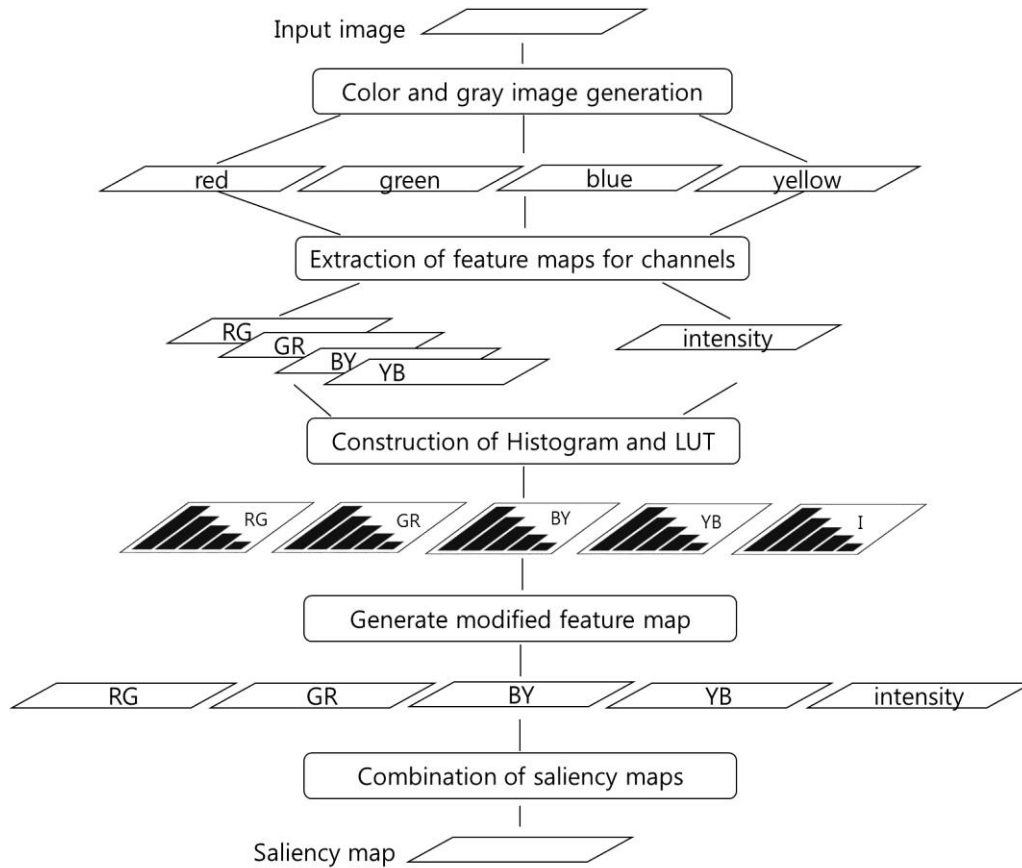


Figure 1. Block Diagram of Proposed Algorithm

The first step is to generate yellow component and intensity values of pixels from an RGB color image. From color components, it generates contrast values between opponent colors at the second step. Considering the contrast values and the intensity values as separate channels, the algorithm constructs histograms of values for each channel and the histograms are normalized according to their relative frequencies. At the final step, the intended saliency map is generated by converting and combining the histograms.

3.1 Generation of Yellow and Intensity Maps

From an input RGB color image, we build four channels of color component by generating yellow component in addition to R, G and B components. We also build a channel of image containing gray scale intensity values. The reason for having four components of colors in this algorithm is that human perception of colors is based on the opponent colors [6]. Yellow components and intensity values at each pixel locations are generated through a simple calculation of $(\text{Red} + \text{Green}) / 2$ and from $(\text{Red} + \text{Green} + \text{Blue}) / 3$, respectively.

3.2 Extraction of Color Feature Images

In human color perception process, a person recognizes a variety of colors using contrast between pairs of "red-green" and "blue-yellow". According to the theory, in this paper, features using difference of Red-Green, Green-Red, Blue-Yellow and Yellow-Blue, as shown in Eq. 1, are used for finding salient objects.

$$\begin{aligned}
 RG &= \langle \text{Red-Green} \rangle \\
 GR &= \langle \text{Green-Red} \rangle \\
 BY &= \langle \text{Blue-Yellow} \rangle \\
 YB &= \langle \text{Yellow-Blue} \rangle
 \end{aligned}
 \tag{1}$$

'-' represents the pixel-by-pixel difference between the color components, and when a value becomes negative, it is set to zero.

3.3 Generation of Histogram and LUT

With RG, GR, BY, YB and I feature channels, histograms of these five channels are separately constructed by counting the frequency of specific values of each channel at each pixel location. Since more salient objects perceived by human observers tend to take relatively smaller area in an image, the histograms are modified and normalized considering the tendency. The modification of each histogram are implemented as a set of look-up Tables (LUT). Eq. 2 shows the modification process.

$$\begin{aligned}
 N &= 256 \\
 S_i &= \sum_{j=0}^{N-1} |i-j| H_j, (i = 0 \dots 255)
 \end{aligned}
 \tag{2}$$

In this equation, i and j are indices for strengths of a channel and H_j is the corresponding frequency in a histogram. The equation collectively reconstruct a histogram as a function of distance between two strengths and the frequencies of all other strengths. In another word, " H_j " is a value of j-th bin of a histogram and " S_i " is a cumulative sum value of histogram using weight of distance between i-th bin and all other bins. If a cumulative sum value becomes large, it may be inferred to a chance of having high saliency value.

The histograms are further modified for normalization purpose using the following equations.

$$NH_i = \frac{S_i - \min(S)}{\max(S)}, (i = 0 \dots 255)
 \tag{3}$$

$$G_i = 255(NH_i)^{-\gamma}, (i = 0 \dots 255)
 \tag{4}$$

In Eq.3, NH represents the normalized histogram. The final histograms used in this paper is further modified using the gamma transformation as in Eq.4 the best value experimentally found for the gamma was 1.2. These five channels with modified histograms (or lookup tables) are used to reconstruct feature maps for each channel.

3.4 Conversion of Histograms to a Saliency Map

Combining the evidences for saliency map from five channels is done by average operation. We firstly change the original color and intensity channels using lookup tables by assigning new channel values based on the values in the corresponding lookup table. The final saliency map is obtained by averaging values from five feature maps at the same pixel locations.

4. Experiment and Analysis

We compare our proposed method with three previous methods including two methods introduced in Section 2 using ROC analysis. The database used in this paper is IVRG data set [7], which contains one thousand labeled still images. The ground truth for salient objects are given as binary images as shown in the last column in Figure. 2. For ROC

analysis, two rates, sensitivity (True Positive Rate) and specificity (false negative rate), are calculated as shown in Eqs. 5 and 6 [8].

$$TPR = \frac{\text{True Positive}}{\text{True Positive} + \text{True Negative}} \quad (5)$$

$$FPR = \frac{\text{False Positive}}{\text{False Positive} + \text{True Negative}} \quad (6)$$

The final saliency map from the proposed method are provided as gray-level images. To obtain ROC curves, the resulting saliency maps are thresholded with different values starting from 0 to 255, and compare with the ground truth images.

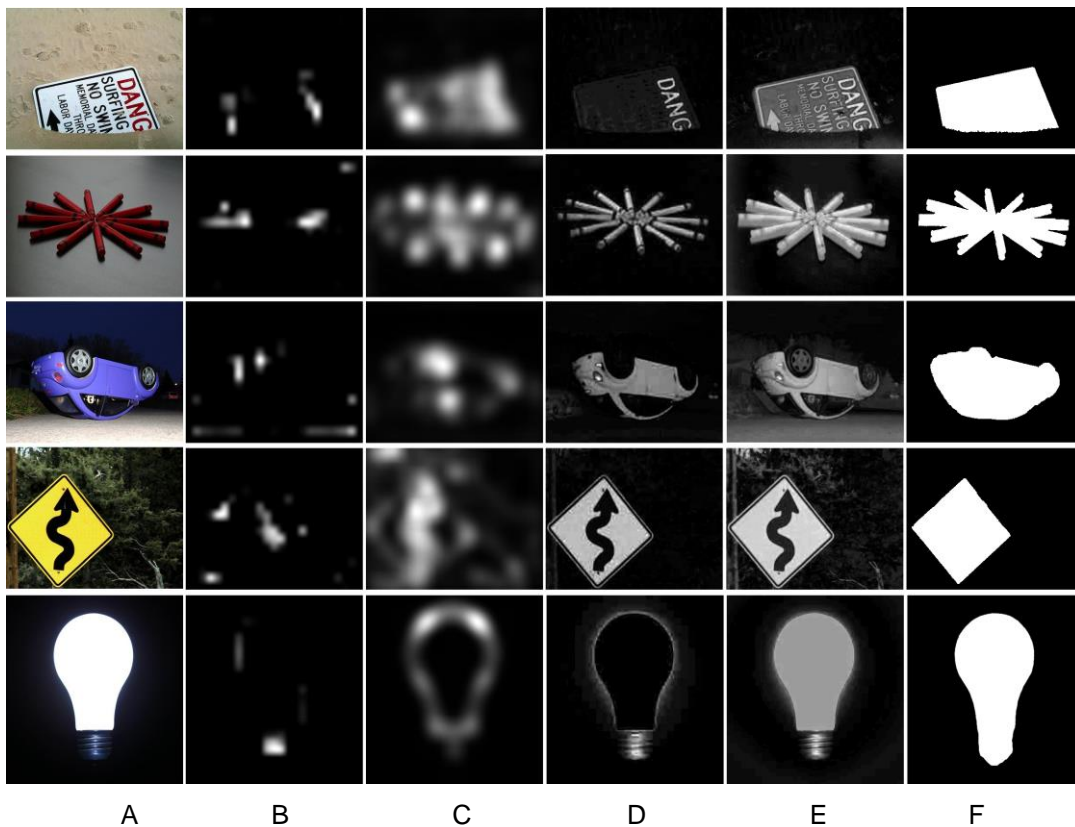


Figure 2. From Left A. Original Image B. IT Result Image C. SR Result Image D. Only Use Color Method Result Image E. Proposed Method Result Image F. Ground Truth Image

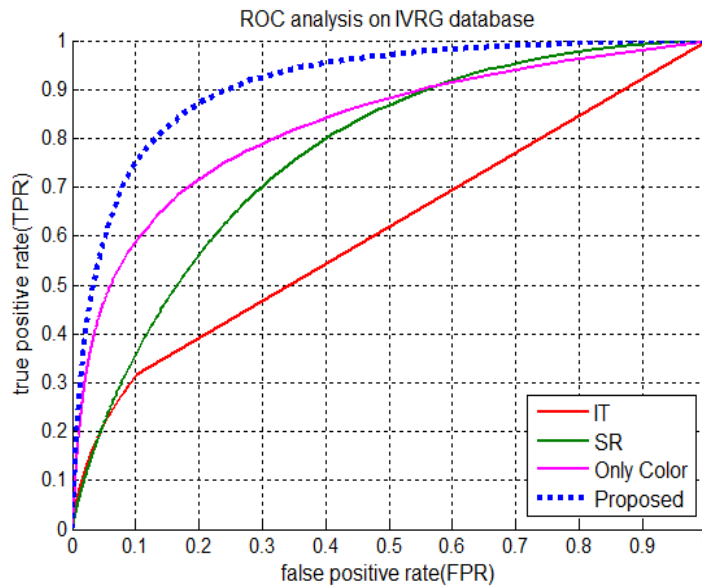


Figure 3. ROC Curve

4.1 Experiment

Figure. 2 shows the experimental results of the proposed method. In the Figure, starting from the leftmost column, it shows five examples of the original input image, saliency maps for IT and SR methods, saliency maps using only color Tables [9], and the results of the proposed method. The last column of the figure shows the ground truth images provided in the database. As we can see in the figure, the proposed method clearly generates saliency maps comparing to the existing methods.

Figure. 3 shows the ROC curves for all the images in the database and it is obvious that the proposed method outperforms the others.

4.2 Analysis

To compare the overall performance of the proposed method with previous methods, we used the AUC values for IT, SR, color channel-only method and the proposed method are 0.60, 0.65, 0.82 and 0.91, respectively. Thus, it can be seen that our method improves detecting performance from about 10 to 40%.

As in Figure. 2, the proposed method accurately describes boundaries of salient objects and their position comparing to other methods. Especially, the proper use of color and brightness information together generally represent salient regions more accurately for images with large contrast than the method of using only color information. However, the additional use of intensity values has introduced small noises if an image has very complex background, as we can see in some images in Figure.2.

5. Conclusions

Generating saliency maps is a method to find the area of interest in the process of image understanding. In this paper, we proposed a detector to improve disadvantages of IT and SR methods. Generally, The IT method tends to not clearly describe whole salient objects, and SR method does not clearly detect position of salient object. The proposed detector is based on four color components, red, green, blue and yellow, and brightness information. We used the concept of opponent colors to construct the color feature maps and their histograms. The histograms are modified and normalized to reconstruct the color channels representing modified

feature maps. The actual implementation is done by using a set of lookup tables for this purpose. Finally, the saliency map is obtained from modified LUT.

The proposed detector outperforms other existing methods by increasing the AUC about 10-40%. Besides, the resulting salient objects show very clear boundaries with the very close estimation of positions of salient objects location.

Acknowledgements

This work (Grants No. C0213585) was supported by Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration in 2014.

This work was supported by the Brain Korea 21 PLUS Project, National Research Foundation of Korea.

References

- [1] Sang B. P., Ki J. K., Young J. H. and Hern S. H., "Implementation of Image Adaptive Map," Journal of the Korean Society for Precision Engineering, vol. 25 no. 2, (2008), pp. 131-139.
- [2] Laurent I., Christof K. and Ernst N., "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20 no. 11, (1998), pp. 1254-1259.
- [3] Xiaodi H. and Liqing Z., "Saliency Detection: A Spectral Residual Approach," Computer Vision and Pattern Recognition, (2007), pp.1-8.
- [4] A. Oliva, A. Torralba, M. S. Castelhano and J. M. Henderson, "Top-down control of visual attention in object detection," International Conference on Image Processing, vol.1, (2003), pp. 253-256.
- [5] M. Milanova, S. Rubin, R. Kountchev, V. Todorov and R. Kountcheva, "Combined visual attention model for video sequences," International Conference on Image Processing, (2008), pp. 1-4.
- [6] Te W. L., Thomas W. and Terrence J. S., "Color opponency is an efficient representation of spectral properties in natural scenes," Vision Research, vol. 42, (2002), pp. 2095-2103.
- [7] "http://ivrgwww.epfl.ch/supplementary_material/RK_CVPR09/index.html."
- [8] "http://en.wikipedia.org/wiki/Receiver_operating_characteristic."
- [9] Changpyo H., Kiseon J. and Dong S. P., "Design and Implementation of Efficient Saliency Detector using Color LUT", Proceedings of KISM Fall Conference 2013, vol. 2 no. 2, (2013), pp. 116-119.

Authors



Changpyo Hong, He received the B.S degree from Chonbuk National University, Jeonbuk, Korea, in 2014. Now he is M.S student with Chonbuk National University, Korea. His current research interests include image processing.



Jong Bin Park, He received the M.S degree from Chonbuk National University, Jeonbuk, Korea, in 2004. Now he is Ph.D. student with Chonbuk National University, Korea. His current research interests include machine learning and deep belief network.



Sook Yoon, She received the B.S., M.S., and Ph.D. degrees in engineering from Chonbuk National University, Jeonbuk, Korea, in 1993, 1995, and 2003, respectively. Until June 2006, she conducted her postdoctoral research work in electrical engineering at the University of California, Berkeley. She is presently an associate professor at Department of Multimedia Engineering, Mokpo National University, Jeonnam, Korea. Her current research interests include image processing, pattern recognition, machine learning, and multimedia computing.



Dong Sun Park, He received the B.S. degree from Korea University, Seoul, Korea, in 1979, and the M.S. and Ph.D. degrees from the University of Missouri, Columbia, in 1984 and 1990, respectively. He is currently a Professor with the School of Electronic Engineering, Chonbuk National University, Jeonbuk, Korea. His current research interests include image processing, pattern recognition, computer vision, and artificial intelligence.