

# Saliency Region Detection via Local and Global Features

Dan Hu

School of Electronic and Information Engineering  
South China University of Technology  
Guangzhou, P. R. China

Yan Wang

School of Electronic and Information Engineering  
South China University of Technology  
Guangzhou, P. R. China

Shaohui Qian

School of Electronic and Information Engineering  
South China University of Technology  
Guangzhou, P. R. China

Wei Yu

School of Electronic and Information Engineering  
South China University of Technology  
Guangzhou, P. R. China  
yuweiyu@scut.edu.cn

**Abstract**—Saliency detection plays an important role in many applications such as image segmentation, retrieval, editor and object recognition. Saliency detection methods mainly divide into two categories: local contrast method and global contrast method. AC method is one of the typically local contrast methods, and HC method is a classical method base on global contrast. In this paper, we proposed an approach combined with AC and HC method, it not only considers local features but global information. After synthesizing the saliency map of the AC and HC methods, firstly, according to the characteristics of human visual system, we pay attention to the saliency center of an image, and add a weight value to each pixel, the closer to the center, the higher the weight. Then, we enhance the contrast of the image, receding darker part and increasing bright one. To confirm the proposed approach's effectiveness, we compare AC and HC method with the proposed one. The experimental results show that the new method generates a better saliency map.

**Keywords**—saliency detection; local contrast; global contrast; weight value

## I. INTRODUCTION

Human visual system can rapidly detect salient region of natural and synthesis image for further high-level cognitive processing. Due to the importance of saliency detection for reducing computational complexity, modeling saliency detection has attracted numerous research efforts. People always tend to find a saliency map and help us getting important information in a scene, so it becomes a popular topic of saliency detection. Saliency detection have many application in image processing and computer vision, such as image segmentation[1], image retrieval[2], object recognition[9], image compression[10], image quality assessment[11,12], and so on. Based on visual attention mechanism, saliency detection methods fall into two categories, bottom-up and top-down method. The bottom-up approach mainly combined with center-neighborhood

structure to achieve a saliency molding. There are 3 steps in our method: first, extract the initial characteristics of image; second, calculate saliency of local region through center-neighborhood differences; final, merge all characteristics to a final saliency map. Unlike the bottom-up one, top-down approach integrates prior knowledge of statistics feature to get the salient map. Most of the saliency detection methods belong to the former one: bottom-up model. After years of development, many methods have been proposed to obtain saliency map. Itti et al. [3] introduced a model to calculate center-surround contrast base on biological structure. The model used 9-scale image features which contain color, intensity, orientation and other characteristics. Zhai et al. [4] consider each pixel's contrast in the entire image as saliency value. Hou et al. [5] directly analyze logarithmic spectrum of the image. R. Achanta et al. [6] presented a method to determine salient regions from the frequency perspective. R. Achanta et al. [7] also use luminance and color features of its neighborhood at various scales to estimate salient region. Cheng et al. [8] extract the saliency map based on global histogram contrast. And they also present a method (RC) to divide the image into blocks, and then count contrast of each block.

Detection algorithms can be classified into local contrast method and global contrast according to contrast type. As mentioned above, IT, MZ, CA, AC are local contrast methods while HC, LC, RC, SR, FT are global contrast method. We propose a new algorithm combined with local contrast and global features.

The rest of this paper is organized as follows: Section 2 local saliency is introduced. In Section 3, Global Salient Detection is described. In section 4, a new approach is proposed. Experimental results are presented in Section 5, while Section 6 discusses the conclusions.

## II. LOCAL SALIENCY DETECTION (AC)

AC method is a typical local contrast based saliency detection approach. It calculates local contrast in neighborhood of each perceivable unit, every unit at different scales. The step is as follows:

First, divide the input image into perceivable units which can be a sub-region. The pending unit is inner region  $R_1$  and the outer region  $R_2$  is a template surrounding  $R_1$ .  $L$  is the size of the image.

Second, keep the  $R_1$  and change the size of  $R_2$  from  $\frac{L}{8}$  to  $\frac{L}{2}$ . Calculate the region saliency value of  $R_1$  in every  $R_2$ 's scale, then have the average of these saliency values as final saliency value of  $R_1$ .

Third, change the next unit as  $R_1$  and repeat step 2 until all of the units have been processed.

It is usually given 3 sizes of  $R_2$  ( $\frac{L}{8}$ ,  $\frac{L}{4}$  and  $\frac{L}{2}$ ). At a given scale, the saliency value of  $R_1$  is determined by the distance between the average vectors of  $R_1$  and  $R_2$ , the vector contains feature of each pixel:

$$S(i, j) = d\left[\left(\frac{1}{N_1} \sum_{k \in R_1} v_k\right), \left(\frac{1}{N_2} \sum_{k \in R_2} v_k\right)\right] \quad (1)$$

$S(i, j)$  is saliency value of  $R_1$  at position  $(i, j)$  in the image.  $N_1$  and  $N_2$  mean the number of pixels in  $R_1$ ,  $R_2$ ,  $v$  is vector of feature value to a pixel.  $d$  is a Euclidean distance of  $v$  in the CIELab color space. CIELab color contains luminance, color information and it is similar to human vision perception. The feature vector  $v$  is expressed as  $[L, a, b]$  in CIELab color space. Therefore, Equ. (1) can be replaced as bellow:

$$S(i, j) = \|v_1 - v_2\| \quad (2)$$

Where  $v_1 : [L_1, a_1, b_1]$  and  $v_2 : [L_2, a_2, b_2]$  are the average feature vector values of  $R_1$  and  $R_2$ . In Equ. (2), we don't calculate distance between each pixel of  $R_1$  to each pixel of  $R_2$ , instead, we can just compute the average feature vector values of  $R_1$  and  $R_2$ . We get the final feature values of  $R_1$  in the position  $(i, j)$  by integrating the feature of  $R_2$  with each given scale. We can obtain the saliency map by processing every  $R_1$  of the image with the same approach.

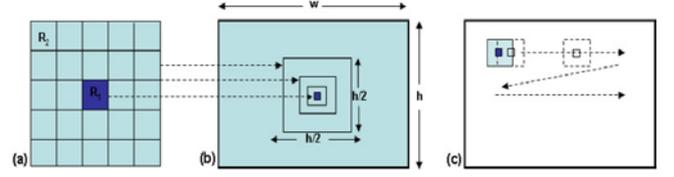


Figure1 (a) Inner square region  $R_1$  and outer square region  $R_2$ . (b)The width of  $R_1$  fixed while width of  $R_2$  ranges from  $h/8$  to  $h/2$ . (c) sketch of processing the whole image.

The process is described in figure1. As shown above, the final feature values of  $R_1$  is integrated by the situation in mutative size of  $R_2$ . We obtain the saliency map by processing every  $R_1$  of the image with the same approach.

## III. GLOBAL SALIENT DETECTION (HC)

In contrast, HC method is a typical saliency detection approach based on global contrast. It is a histogram-based method to measure saliency that calculate color deference with all the other pixels to assign pixel-wise saliency values. Saliency value of pixel  $I_k$  in the image  $I$ ,  $S(I_k)$  is described in equ.(3).

$$S(I_k) = \sum_{\forall I_i \in I} D(I_k, I_i) \quad (3)$$

Where,  $D(I_k, I_i)$  refers to Euclidean distance between pixel  $I_k$  and pixel  $I_i$  in the CIELab color space. Equ. (3) can be written as (4):

$$S(I_k) = D(I_k, I_1) + D(I_k, I_2) + \dots + D(I_k, I_N) \quad (4)$$

$N$  is the number of pixels in image  $I$ . All the pixels with same color have the same saliency values. So the same color pixel can be grouped together and equ. (4) can be written as bellow:

$$S(I_k) = S(c_l) = \sum_{j=1}^n f_j D(c_l, c_j) \quad (5)$$

Where  $n$  is the number of pixel color in image  $I$ ,  $c_l$  is the saliency value of group  $l$  and  $f_j$  is the probability of  $c_j$  in image  $I$ .

## IV. PROPOSED APPROACH

We proposed a new saliency detection approach based on local and global features of image. It considers not only local contrast features but also global contrast information. Here are several steps as follows:

First, synthesize the saliency map of AC and HC methods to a new saliency map  $S_1$ , as shown in equ.(6):

$$S_1 = a * S_{ac} + b * S_{hc} \quad (6)$$

Where  $S_{ac}$  and  $S_{hc}$  are AC and HC saliency map, while  $a$ ,  $b$  are weights of them respectively. The value of  $a$  is 0.65 and  $b$  equals to 0.35 in experiment.

Second, we calculate the salience center coordinates of an image:

$$x = \frac{\sum_{p=1}^n x_p S_p}{\sum_{p=1}^n S_p}, y = \frac{\sum_{p=1}^n y_p S_p}{\sum_{p=1}^n S_p} \quad (7)$$

Where  $(x, y)$  is salience center coordinate of salience map  $S_1$ .  $x_p, y_p$  are horizontal ordinate and vertical coordinate of the pixel  $p$ , whose salience value is  $S_p$ .  $n$  is the total number of pixels in image  $S_1$ . Then add a weight value to each pixel by Euclidean distance from salience center, the closer to the center, the higher the weight.

$$S_2(p) = S_1(p) \times \exp\left(-\frac{D_p}{\varepsilon}\right) \quad (8)$$

Where  $D_p$  is Euclidean distance between pixel  $p$  and the salience center  $(x, y)$ .  $S_1(p)$  is salience value of pixel  $p$  in map  $S_1$ .  $\varepsilon$  is a constant-adjustment coefficient valued 500.

Third, after space processing, we enhance the contrast of the image  $S_2$  by receding darker part and increasing bright one, it is end up the final salience value  $S_{final}(p)$  of pixel  $p$ . The processing is shown in Equ.(9)

$$S_{final}(p) = \begin{cases} k \cdot \exp\{\delta(S_2(p) - v_{mid})\} + v_{mid} - 1, & S_2(p) \geq v_{mid} \\ -k \cdot \exp\{-\delta(S_2(p) - v_{mid})\} + v_{mid} - 1, & S_2(p) < v_{mid} \end{cases} \quad (9)$$

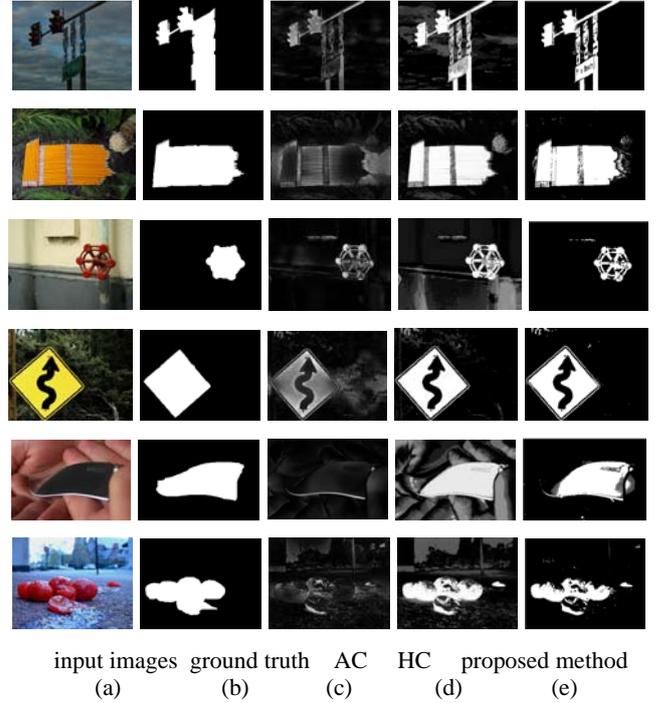
$$v_{mid} = \frac{2}{3}(v_{max} - v_{min}) \quad (10)$$

1) Where  $0 \leq S_{final}(p) \leq 255$ ,  $v_{max}$  and  $v_{min}$  are the maximum and minimal value in image  $S_2$ .  $k, \delta$  are certain coefficients valued 0.5 and 0.6 in experiments.

## V. EXPERIMENTAL RESULTS

In experiment, software is Matlab 2010, 2.0GHz CPU, and 2.0GHz Memory computer. We compared saliency map of our method with that of AC and HC method from MSRA database. Experiment results is listed in figure1, AC method concentrates

on local features but ignores global features, so it can easily detect tiny edge information while the entire image is much dark; As opposed to AC method, the HC method only considers global features and takes on account of local information, therefore, some details haven't been handled very well on the saliency map. It combines peculiarity of AC method and HC method of our proposed method, which pays attention to both local traits and global traits. The new saliency map shows the saliency object more clearly and exactly than AC map or HC map.



## VI. CONCLUSION

An image usually contains both local features and global features, and local or global information is just one aspect of image. It is useful to combine two aspects features of image. In this paper, we proposed a saliency detection approach which uses local features as well as global features of an image. In experiment, we compare proposed approach with the AC method and HC method. Experiments results demonstrated that proposed approach based on comprehensive features achieves better effects than AC or HC method in extracting saliency map.

## VII. ACKNOWLEDGEMENTS

This work is supported by Undergraduate Student Self-select Programs Supported by Fundamental Research Funds for the Central Universities, South China University of Technology. (grant No. 10561201442)

## REFERENCES

- [1] S. Goferman, L. Zelnik-Manor, and A. Tal. Context-aware saliency detection. in IEEE CVPR, 2010, pp. 2376-2383.
- [2] S.-M. Hu, T. Chen, K. Xu, et al. Internet visual media processing: a survey with graphics and vision applications, *The Visual Computer*, 2013, pp. 1-13.
- [3] L. Itti, C. Koch, E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), 1998, 1254-1259.
- [4] Y. Zhai and M. Shah. Visual attention detection in video sequences using spatiotemporal cues. in *ACM Multimedia*, 2006, pp. 815-824.
- [5] X. Hou and L. Zhang. Saliency detection: A spectral residual approach, in *IEEE CVPR*, 2007, pp. 1-8.
- [6] R. Achanta, S. Hemami, F. Estrada, et al. Frequency-tuned salient region detection, in *IEEE CVPR*, 2009, pp. 1597-1604.
- [7] R. Achanta, F. Estrada, P. Wils, et al. Salient region detection and segmentation. In *ICVS*, 2008, pp.66-63.
- [8] M. M. Cheng, G. X. Zhang, N. J. Mitra, et al. Global contrast based salient region detection. In *CVPR*, 2011, pp. 409-416.
- [9] U. Rutishauser, D. Walther, C. Koch, and P. Perona. "Is bottom-up attention useful for object recognition?" in *IEEE CVPR*, 2004.
- [10] C. Christopoulos, A. Skodras, and T. Ebrahimi. The JPEG2000 still image coding system: an overview, *IEEE Trans. Consumer Elec.*, vol. 46, no. 4, 2002, pp. 1103-1127.
- [11] A.K. Moorthy, A.C. Bovik, Visual importance pooling for image quality assessment, *IEEE J. Select. Top. Signal Process.* 3, 2009, 193-201.
- [12] J. You, A. Perkis, M.M. Hannuksela, et al. Perceptual quality assessment based on visual attention analysis, in: *ACM Int'l Conf. Multimedia*, pp. 561-564.