A New Salient Region Detection Based on Biomimetic Information

Zongmin Li
China University of Petroleum(east China)
College of Computer and Communication Engineering
Qingdao, China
lizongmin@upc.edu.cn

Jinfeng Chen
China University of Petroleum (east China)
College of Computer and Communication Engineering
Qingdao, China
chenjinfeng07@sina.com

***Abstract*—In our paper, we develop a new method of detecting salient region. It outputs full resolution saliency maps with uniformly highlighted the entire salient region. In our proposed method , we incorporate the traditional low-level features with high dimension biomimetic information geometry theory and priori knowledge. Extensive experiments on one of the largest public salient object detection benchmark have demonstrated that our method can effectively achieve good performance to the exiting methods. Our method has obtained higher precision and better recall rates on the segmentation task comparing to the five advanced salient region detection algorithms.**

Keywords-salient object detection, biomimetic information, image features, multi-scale saliency

I. Introduction

The digital image increases dramatically, it is particularly important to obtain the useful information and remove useless information accurately of the image at the same time. Salient region detection is one of the methods to solve this problem. In nowadays, salient detection has attracted more attention and used widely in a lot of applications ,for example, adaptive compression [1], image segmentation[2], object recognition or content-based image retrieval [3]. More and more saliency region detection algorithms are proposed. In general, the saliency estimation methods are divided into three categories. The first is on account of low-level vision features of images, for example, the color feature and the texture. The second is based on purely computational. The third is based on a combination of the two methods.

Salient region detection is closely related with the biological visual theory. Inspired by the biomimetic information geometry theory, we present a new saliency region algorithm, which can able to detect the visually salient regions in images automatically and effectively. First of all, we segment the input image into two regions, one is the main region and the orther is the margin region which is based on the biomimetic information geometry theory. Secondly, we compute average vector of margin region of the input image using color and illumination in the CIELab color space. Then, we calculate the Euclidean distance of the average CIELab vector of margin region with each pixel of a Gaussian blurred version (using 5*5 binomial kernel) as the saliency value.Besides, we use the Gaussian weight to revise the saliency. The experiment results on the number of a 1000 images datebase demonstrate that we have achieved outstanding performance according to precision and recall against five state-of-the-art methods. Some saliency maps which got by our method are shown in figure 1.

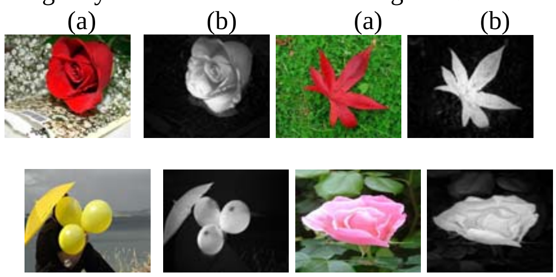


Figure 1. Some example of saliency maps detected by our method,(a) are the original images. (b) are our's saliency map images.

II. Review Of The Related Methods

Salient region has the property is that it is highlight to their surroundings and is easily to attract human attention. In recent years, the large number of saliency detection methods were developed. The well known biological method is put forward by Itti[5] ,which is on account of the principle put forward by Koch and Ullman. They exploit the difference of gaussians (DoG) method tocompute the center surround contrast .

Hou and Zhang [6] put forward the spectral residual hypothesis and extracted the input image's spectral residual to construct the homologous saliency map in spatial field. In [7], they proposed an approach based on contrast analysis, which generated the saliency map by using fuzzy growing. In [8], they calculate saliency using center-surround feature distances. Achanta [4] proposed a frequency-tuned algorithm which obtain the saliency map through computing the Euclidean distance which is the average CIELab vector of all pixels with every pixel of a Gaussian blurred version of the same input image. Subsequently, in[9], Achanta improved the frequency-tuned algorithm, which also exploited features of color and luminance. Instead, they computed the average distance which is the CIELab vector of a symmetric surround region for the pixel with the pixel instead. Cheng[10] proposed the RC method that incorporated regional and global contrast-based and features spatial relations. The region's saliency value is computed by making use of the global contrast-based score, measured by the contrast of the region and spatial distances to other region's in the image. To some degree, dividing the image into multiple regions reduced calculation efficiency.

There are many methods based on incorporating the biological models with the computational models. For example, in [11] they proposed a visual saliency model that formed activation maps on some feature channels, then normalized them.

III. Overview Of Our Method

In our paper, we present a new algorithm which is used for detecting saliency region. It makes use of biomimetic information geometry theory and low-level features . Firstly, we get three scales through Gaussian Pyramid. For each scale, we devide the image into main region and margin region which is based on high dimension biomimetic information geometry theory. Secondly, computing the average CIELab vector of pixels for margin region. To avoid the impact of image noise , we smooth the input image using the Gaussian filter . Then, the saliency value of each pixel is determined by the Euclidean distance to the average vector of margin region. In the end, we obtain the Gaussian weight using Euclidean distance from each pixel to the center of the image .The framework of our algorithm is shown in Figure 2.

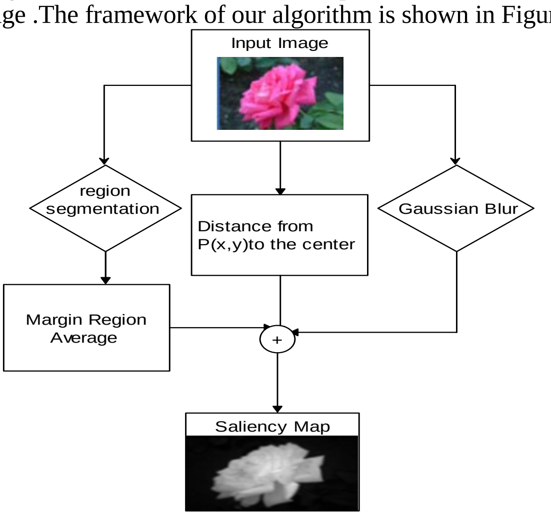


Figure 2. The flawchart of our method.

A. *Color Space Conversation*

It is important to pay attention that the color space should be accord with human color vision system when extract the color feature of the image. At present, the most commonly color space is the RGB in computer vision. However, it isn't accord with human color perception system well. Many researches have found the Lab color space is well accord with human vision system. Therefore, we employ the Lab color space. As many images that we obtain are RGB images , what's more Lab color space and RGB color space cannot convert each other directly, in this case, with the aid of XYZ color space , image is transformed from RGB space to Lab space through the function as follows:

$$L = 116 \times f(Y/Y_n) - 16 \tag{1}$$

$$a = 500 \times \left[f(X/X_n) - f(Y/Y_n) \right] \tag{2}$$

$$b = 200 \times \left[f(Y/Y_n) - f(Z/Z_n) \right] \tag{3}$$

f is correction function which is similar to the Gamma function, specifically:

$$f(t) = \begin{cases} t^{1/3}, & \text{if } t > (6/29)^3 \\ \frac{1}{3} \times (\frac{29}{6})^2 \times t + 16/116, & \text{if } t \le (6/29)^3 \end{cases} \tag{4}$$

n denotes the value has been normalized, and the relationship between XYZ and RGB is defined as:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \bullet \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{5}$$

B. *Region Segmentation*

The high dimension biomimetic information geometry theory has been widely used in many researches [12] . As far as we know, this theory has not been used to saliency region detection. Based on this theory, different from the method of [10], we segment the input image into two regions,one is the main region and the other is the margin region. In our paper, the width of margin is 0.25 of image size. In paper [13], they selected the width of margin as 0.1 of image size. In general, the object is often located in the center region of image, so the main region which stands the main object . The margin region provide the surroundings of the object, it also plays a helpful role in saliency region detection. So we can combine the main region with margin region to detect the saliency region..

C. *Computing Saliency Region*

We compute the center-surround contrast making use of color and luminance features. In common, all of methods put to use a low-level method by computing contrast of image regions compare to their surroundings, using one or more features,such as the intensity, color, and orientation.

In[4] their saliency map is achieved through computing the Euclidean distance of the average CIELAB vector of all pixels of an input image with each pixel of a Gaussian blurred version of the input image. However, when the area of the salient region is large, or the background is mazy, the saliency maps stand out the background. In order to overcome this drawback, we computer the average CIELab vector of pixels of margin region instead of computing the the average CIELab vector of pixels of the whole image. Then, we calculate saliency map S for an image as follows Eq(6).

$$Salience(x, y) = \left\| I_{evu} - I_{gv}(x, y) \right\| \tag{6}$$

Where Ievu is the arithmetic mean pixel value of the margin region. Igv (x, y) indicates the Gaussian blurred version of the original image at the point of (x, y). || || is the L2 norm..

Considering the people tend to stare at the center of the image according to the previous studies on the distribution of human fixations on images [14]. What's more, people always put the target in the center or near the center of the image when they take pictures. In [15] a method which center bias had been taken into account and shown to improve the performance significantly. In [16], they explore saliency region detection approach which involving center prior. Inspired by above methods, the distance of each pixel from the center of the image is involved in our evaluation of the saliency map. So we rewrite Eq(6) as:

$$Saliency(x,y) = w((x,y),center) * \left\| I_{evu} - I_{gv}(x,y) \right\| \quad (7)$$

Where , w(x,y) is derived from the Gaussian function which defined as:

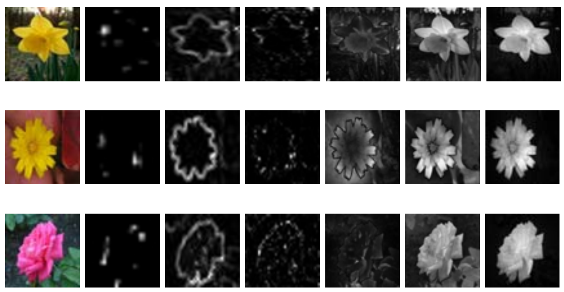
$$w(x,y) = e^{-d^2(x,y)/2\sigma^2} \quad (8)$$

d(x,y) is the spatial distance between the point of (x,y) and the center point of the input image. σ is the longer side of the input image.

IV. EXPERIMENT AND ANALYSIS

In this part, we verified our method on the one of the largest public data sets which provided by Achanta [4]. The data set contains 1000 images. To save the space, we only compare our method with five advanced methods, namely IT[5], MZ[7], SR[6], AC[8], FT[4]. In Figure 3, we compare the saliency maps obtained by the different methods. As can be seen, our method achieves the best visual saliency maps.

As is known to all , saliency detection has been widely used in object segmentation . So as to evaluate the accuracy of our methods for salient object segmentation, we have done two experiments. In our first experiment, we segmented the object by fixed thresholding. In another experiment, we segmented the object by adaptive adaptive thresholding.



(I) (a) IT (b) MZ (c) SR (d) AC (e) FT (f)Our's

Figure 3. Visual comparision of saliency maps(I) are the original images.(a) are the saliency images of IT[5] (b) are the saliency maps of MZ[7] (c) are the saliency maps of SR[6] (d) are the saliency maps of AC[8] (d) are the saliency maps of AC[8] (e) are the saliency maps of FT[4] (f) are the saliency maps of our method.

As is known to all , saliency detection has been widely used in object segmentation . So as to verificate the accuracy of our methods for object segmentation, we have done two experiments. In our first experiment, we segmented the object by fixed thresholding. In another experiment, we segmented the object by adaptive adaptive thresholding.

A. Segmentation by Fixed Thresholding

Once we have obtained a saliency map , normalized it between [0, 255], a binary mask is generated by thresholding the saliency map at each threshold ranging from 0 to 255. So there will be 256 different binary masks for a saliency map that has values ranging from 0 to 255. Next we compare each such binary map with the ground truth , which is also a binary map(values greater than zero indicate object, and values equal to zero indicate background). To compare the effectiveness of the different saliency maps, the value of the threshold ranging from 0 to 255, and calculate the precision and the recall at evry value of the threshold. The precision and recall are defined as follows :

$$\Pr esicion = \frac{\sum S_g \cap S_d}{\sum S_d} \quad (9)$$

Where, Sg represents the saliency region in ground truth , while ,the sd represents the saliency region which detected by saliency detection method. The result of precision-recall curve is shown in Fig.4.

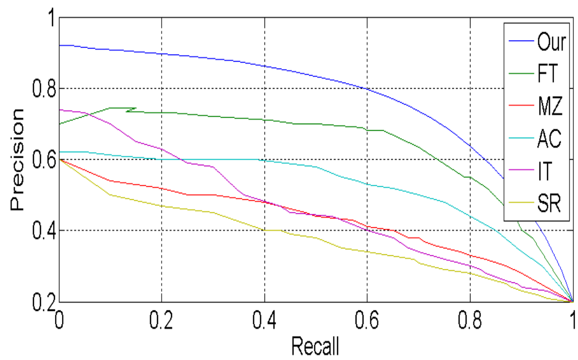


Figure 4. The Precision and the Recall curve for fixed thresholding of saliency map . We compared with five methods of FT[4], IT[5], SR[6], MZ[7], AC[8]

In another kind of evaluation method, we used the adaptive thresholding to generate a binary mask. Average precision, recall, and F -Measure are compared over the same database , The formulation of F-Measure defined as:

$$F_\beta = \frac{(1+\beta^2)\,precision \times recall}{\beta^2 \times precision + recall} \quad (11)$$

In our paper ,we set the β^2 to 0.3 . The comparis on Precision-Recall bars of six methods for saliency maps is shown in Fig. 5.

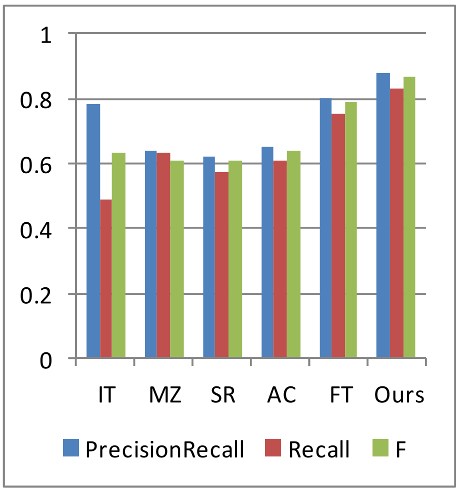


Figure 5. The Precision and the Recall bars for saliency maps.The method of ours show high Our method show high precision, recall, and F-values.

V. CONCLUSION

In our paper, we have come up with a new saliency region detection method which is computationally efficient. We first use the high dimension biomimetic information to compute saliency region . Making use of the "things cognition" instead of "things partition" theory , we devide the image into two regions. Considering the people like to stare at the center of the image according to the studies and always put the target in the center or near the middle of the image when they take pictures, we combine center prior in our method. Also,we combine multi-scale to compute the saliency maps.

In the future, we will further expand our method by incorporating other features to meet different applications.

ACKNOWLEDGMENT

This work is partly supported by the Scientific Research Foundation for the Excellent Middle-Aged and Youth Scientists of Shandong Province of China (Grant No.BS2010DX037), the Shandong Provincial Natural Science Foundation(Grant No.ZR2009GL014) , the Open Project Program of the State Key Lab of CAD&CG (Grant No. A1315), Zhejiang University ,the Fundamental Research Funds for the Central Universities(Grant No. 10CX04043A, 10CX04014B, 11CX04053A, 11CX06086A, 12CX06083A, 12CX06086A, 13CX06007A, 14CX06010A, 14CX06012A).

REFERENCES

[1] C. Christopoulos , A. Skodras, A. Koike, and T. Ebrahimi. The JPEG2000 still image coding system: an overview. IEEE Transactions on Consumer Electronics, 2000, 46(4):1103–1127

[2] J. Han, K. Ngan, M. Li, and H. Zhang. Unsupervised extraction of visual attention objects in color images. IEEE Transactions on Circuits and Systems for Video Technology, 16(1):141–145, 2006.

[3] Stejic Z, Takama Y, Hirota K. Relevance feedback-based image retrieval interfacein corporating region and feature saliency patterns as visualizable image similarity criteria [J]. IEEE Transactions on Industrial Electronics, 2003,50(5):839-852.

[4] R. Achanta, S. Hemami, F. Estrada, and S. S¨usstrunk. Frequency-tuned salient region detection. In CVPR, 2009.

[5] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11):1254–1259, 1998.

[6] X. Hou and L. Zhang. Saliency detection: A spectral residual approach. CVPR, 2007.

[7] Y.F. Ma and H. J. Zhang. Contrast-based image attention analysis by using fuzzy growing. In ACM International Conference on Multimedia, 2003

[8] R. Achanta, F. Estrada, P. Wils, and S. S¨usstrunk. Salient region detection and segmentation. International Conference on Computer Vision Systems, 2008.

[9] Achanta R, Susstrunk S. Saliency detection using maximum symmetric surround [J]. In ICIP, Hong Kong, China, 2010, pages 2653—2656.

[10] M. M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu. Global contrast based salient region detection. In CVPR, 2011.

[11] J. Harel, C. Koch, and P. Perona. Graph-based visual saliency. Advances in Neural Information Processing Systems, 2007.

[12] S. J. Wang, J. Xu, X. B. Wang, and H. Qin, "Multi-camera humanface personal identification system based on the biomimetic pattern recognition," ACTA ELECTRONICA SINICA, vol.31, no.1, 2003, pp. 1–3.

[13] D. Zhuang, S.J. Wang. Content-based image retrieval based on integrating region segmentation and relevance feedback. ICMT, 2010.

[14] B. W. Tatler. The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. Journal of Vision, 2007.

[15] L.J Duan, C. P Wu, J Miao. Visual Saliency Detection by Spatially Weighted Dissimilarity [C].]In CVPR , Colorado Springs, USA, 2011, pages 473-480.

[16] Goferman S, Zelnik-Manor L, Tal A. Context-Aware Saliency Detection [J]. IEEE TPAMI, 2011, PP(99): 1-1