ISSN (e): 2250 – 3005 || Volume, 09 || Issue, 1|| January – 2019 ||
International Journal of Computational Engineering Research (IJCER)

Improved extraction approach of salient object based on visual attention mechanism

Yongmei Zhang[1,*], Juxia Li[2]
[1]Lecturer,College of Information Science and Engineering,Shanxi Agricultural University, Taigu –030801, Shanxi,China.
[2]Associate professor,Shanxi Agricultural University, Taigu –030801, Shanxi,China.
Corresponding Author: Yongmei Zhang

ABSTRACT: *This paper proposed a novel method for salient object extraction based on visual attention. We extract the saliency map by modified Itti's model. A new optimized method to impair the saliency of border area by a location saliency map is propounded. Meanwhile, the input image was segmented by k-means algorithm combining with Hill-climbing algorithm which can give cluster number automatically. The salient object can be obtained by calculating the average saliency of each segmentation area. Experimental results demonstrated that the proposed method is validity and superiority. The results indicate that the salient object is more accurate by the proposed method.*
KEYWORDS:*Visual attention mechanism, salient object,extraction,Itti model, interesting region.*
AMS Subject Classification: 05C12.

Date of Submission: 03-02-2019 Date of acceptance: 20-02-2019

I. INTRODUCTION

Detecting the salient object in an image is the essential part in a wide range of computer vision researches, e.g. image recognition [1], image quality assessment [2] and image compression [3]. Visual attention mechanism has been studied to detect the object, which can imitate how the behaviors transfer their vision when human viewing a scene.

There are many visual attention models for extracting an object or a region from an input image [4-6]. The most widely used model is Itti's model [7]. A brain-like model was introduced and three types of features are utilized to produce a saliency map. This method consists of three stages. In the first stage, Gaussian pyramid is applied to extract the visual features, such as intensity, color and orientation. In the second stage, a center-surround operation is used to calculate the three conspicuity maps. The last stage, a saliency map is combined by all conspicuity maps linearly. Itti's model has been commonly approved by researchers and widely employed in machine vision problems, including automatic target detection [8], object recognition [9], adaptive image display on small devices [10], image compression and coding [11], and irregularities detection [12]. However, they have not been able to accurately detect where visual attention should be by this model.

In this paper, we proposed an approach to extract the exact salient object from an image. Firstly, extracting more exact saliency map by improved Itti's model and optimizing it by proposed location factor; secondly, segmenting the input image by K-means cluster combing with Hill-climbing algorithm; the last, obtaining the salient object by calculating the average saliency of each segmentation area, the max saliency is the salient object. In section 2, the proposed approach is explicated in detail, including modeling the saliency map, segmenting the image and obtaining the salient object. The experimental results are showed in section 3. In section 4, conclusions are drawn.

II. THE PROPOSED APPROACH

The total edge monophonic set M in a connected graph G is called a minimal total edge monophonic set if no proper subset of M is a total edge monophonic set of G. The upper total edge monophonic number $em_t^+(G)$ is the maximum cardinality of a minimal total edge monophonic set of G.

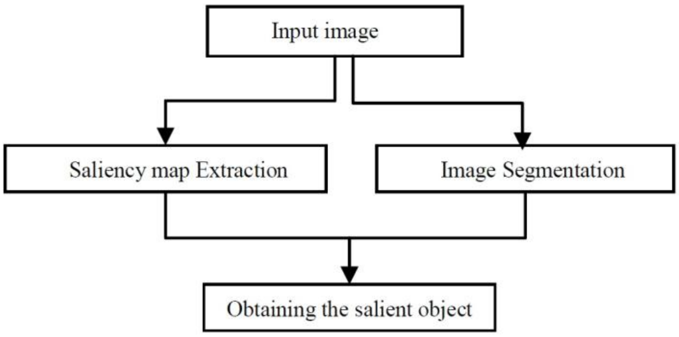

Figure 1.The framework of proposed approach.

Our proposed salient object extraction overall work flow is shown in Fig. 1. The framework considers the saliency regions of an image on segmentation levels and operates in two stages. On the one hand, the saliency map is extracted by improved Itti's model and location factor; on the other hand, the input image is segmented by k-means and Hill-climbing algorithm. Then the average saliency of each segmented area is calculated and the max saliency area is considered as the salient object.

2.1 Saliency map extraction

One of the main goals of visual attention mechanism is to search the most conspicuous parts in the visual field instinctively, which is achieved by estimating the importance of regions at every location. The Itti's model is a classical visual attention model to extract a probability of a region, which represents its importance, is the visual saliency region.

We proposed an improved model to extract more real saliency region based on Itti's model in previous work. The improved Itt's model is described in Ref. [13]. Firstly, three early visual features, such as color, intensity and orientation, were extracted from the input image at multiple scales as Itti's model described. Secondly, three respective conspicuity maps were created according the features by iterative localized interactions combination strategy. The last, in the improved Itti's model, all conspicuity maps were combined into a saliency map nonlinearly. It was considered by Itti et al. that the conspicuity maps with competing strongly from the same features contribute independently and equally to the saliency map. Thus, the final saliency map could be generated by the sum of the three conspicuity maps averagely [7]. In fact, different conspicuity map's contribution to the saliency map should not be equal. We suppose that there is small saliency area in the intensity conspicuity map, which will play a key role in the finally saliency map. Conversely, if the saliency areas distribute uniformly and widely in the orientation conspicuity map, it means that the orientation is not prominent. Evidently, it's contribution to the saliency map is very small.

(a) (b) (c) (d) (e) (f)

Figure 2. the comparison results: (a) original image; (b) color conspicuity map; (c) intensity conspicuity map; (d) orientation conspicuity map; (e) the saliency map by Itti's model; (f) the saliency map by our improved Itti's model.

According to the above discussion, it can be thought that the contribution rate is related to saliency points' area and distribution in some conspicuity maps. The contribution for the saliency map increased with decreasing the saliency points' area in the conspicuity map. That is to say, they're both inversely proportional. The saliency map can be defined as following:

$$S = \sum_{C \in \{C_I, C_C, C_O\}} \eta(C) \mathrm{N}(C) \qquad (1)$$

$$\eta(C_i) = \frac{1/s(C_i)}{\sum_{i \in \{I,C,O\}} (1/s(C_i))} \quad i \in \{I, C, O\} \qquad (2)$$

where C_i is the conspicuity map of intensity, color and orientation. $\eta(C_i)$ is contribution rate of C_i to the saliency map. $s(C_i)$ is saliency points area in C_i .

Fig. 2 shows the comparison result of our improved Itti's model with Itti's model. Fig. 2(a) is an input image. Obviously, the flower is the salient object. Figs. 2(b)-(d) are conspicuity maps. The saliency points' area is distributed widely in the orientation conspicuity, which may result in some saliency points outside the saliency object by Itti's model, as shown in Fig. 2(e). Fig. 2(f) is the saliency map by the improved Itti's model. As shown in this figure, the saliency map is more close to the saliency object. It has been confirmed that the saliency map accuracy increased by 15-20% compared with Itti's model [13].

The eye-tracking experiments have showed that not only the low factors such as contrast, shape, color, motion and orientation can guide human visual attention system but also the high level factors play important parts in visual attention operation including location, context, texture and so on [14]. The experimental results indicated that viewers' eyes are directed at the centre 25% of a screen for a majority of viewing material. Namely, we pay more attention to center of images. Thus, an equation is proposed to calculate each pixel's weight of location as follow:

$$W(x,y)=\begin{cases} 1 & D(x,y)\in D_0 \\ \dfrac{1}{1+\left[D(x,y)/(4\times D_0)\right]^2} & other \end{cases} \quad (3)$$

where $D_0=\sqrt{m\times n}/8$ is the central 1/4×1/4 of the image, m and nare thewidthandheightof the imagerespectively.D is the distance from the pixel (x,y) to the center of the image.

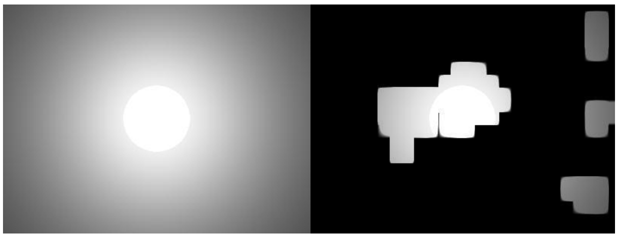

Figure 3.Optimize results: (a) the location saliency map; (b) the optimized saliency map.

Fig. 3(a) shows the location saliency equation (3), which is called as location saliency map in this work. As can be seen, the central 1/4×1/4 area has the max saliency and the saliency decreases with the increase of the distance between the pixel (x,y) and the center. In order to emphasize the saliency of the center area and decrease the marginal saliency, we optimized the saliency map with the location saliency map as follow:

$$SM(x,y)=W(x,y)S(x,y) \quad (4)$$

Where S(x,y) is the saliency map by improved Itti's model, SM(x,y) is the optimized saliency map. The saliency map of Fig. 2(f) is optimized as shown in Fig. 3(b). As shown in this figure,it is easy to observe that the margin area become darker and the saliency become smaller by compared with that of Fig. 2(f).

2.2 Segmentation of image

The saliency object is main area in an image, it has enormous information. An important approach for salient object detection is segmentation. Image segmentation is the process of partitioning an image into non-overlapping regions which allows users to identify meaningful concepts or objects which would be perceptual to human. Cluster analysis is considered as an important measure in the field of pattern recognition. The goal of clustering is to partition data set into such clusters that intra-cluster data are similar and inter-cluster data are dissimilar. In fact, the image segmentation can be regarded as the classification of image pixels. K-means algorithm is suitable for image segmentation and is used widely in image segmentation as the advantages of simple operation and high speed, but different cluster number can lead to different segmentation results. Hill-manipulation algorithm can automatically give cluster number and modes of each cluster [15]. So, it is viable to improve k-means algorithm by hill-manipulation algorithm and segment the image.

Figure 4.Segmentation by the proposed method.

Firstly, the aim image is transformed from RGB to CIE-Lab color space in order that it is most close human vision system; then the number of clustering will be determined by Hill-climbing algorithm. The main idea is that each clustering is expressed by the hill of adjoining colors in the 3D color histogram. The number of hills represents clustering numbers. The value of hill is the initial cluster centers.The distance from pixel to clustering center is defined by square Euclidean distance as followed:

$$D = \{P_i(L^*) - C_n(L^*)\}^2 + \{P_i(a^*) - C_n(a^*)\}^2 + \{P_i(b^*) - C_n(b^*)\}^2 \qquad (5)$$

where$P_i$denotes the simple i, C_n denotes the clustering center n, L^*, a^*, b^* denote the color value in the CIE-Lab color space respectively. Each cluster is represented by the same numbers. The clustering results are shown in Fig. 4.It can be seen that clustering effect is enhanced and the salient object is segmented successfully.

2.3 Salient object extraction

In the optimized saliency map, each pixel's luminance indicates theimportance to human visual attention.In other words, the pixel with higher luminance attracts more attention. According to the segmentation results, we calculate the average saliency of each segmented area when the max saliency segmented area is considered as the salient object. And other small areas in the same cluster with salient object is removed using the erode algorithm of mathematical morphology. At this point, the salient object is extracted exactly.

III. EXPERIMENTS AND DISCUSSION

Our method is applied on the available image dataset, named as MSAR [9], which contains 5000 color images, each image contains at least one salient object. These salient objects differ in color, shape, size, and so on. In other words, there is no more prior knowledge or constraint on these objects except that they are the most salient. So many saliency models for detecting or segmenting salient object evaluate their performance on this image database [16-19].

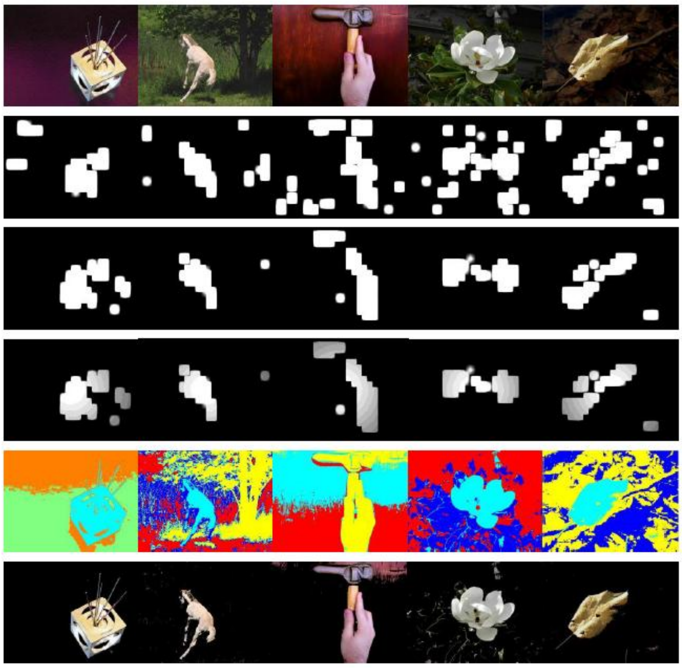

Figure 5.The results by the proposed method. From top to bottom: input image, saliency map by Itti's model, saliency map by our improved model, optimized saliency map by location salient map, segmentation image by clustering combing with Hill-climbing algorithm, the salient object.

Fig. 5 shows the samples of the image dataset. The first row is the input original image; the second row is the saliency map by Itti's model; the third row is the saliency map by our improved model. Comparison of the results with Itti's model suggests that the saliency maps by our proposed model are more close to the saliency objects. The forth row is the optimized saliency map by proposed location salient map.The results show that the intensity becomes darker from the center to the margin. That is, the saliency value becomes smaller and smaller, which indicates that it can simulate attention process when human observe a scene. The fifth row is the cluster results by Hill-climbing algorithm.It shows more clustering effect as the cluster number is unknown. The last row shows that the salient objects are segmented clearly from the background.

IV. CONCLUSIONS

In this paper, we have presented a method to extract the saliency object from the input image without any prior knowledge. More exact saliency map was extracted by improved Itti's model, also we proposed a location saliency model to simulate that observers pay more attention to the center of an image.The salient object was obtained by saliency map based on segmentation. The experiment results show that our method is effective and accurate to extract salient object. It is believed that our approach can be applied to other applications in pattern recognition and computer vision. In future work we will consider other visual factors such as shape, context to extract more exact saliency map.

Acknowledgements
This work issupported by the Key Research Plan Project of Shanxi Province (No. 201803D221028-7).

REFERENCES

[1]. R. Achanta, S. Susstrunk.Saliency detection using maximum symmetric surround. IEEE International Conference on Image Processing. 2010, 2653-2656.
[2]. L. Chen, X.Xie, X. Fan, W. Ma, H. Shang. A visualattention mode for adapting images on small displays. Multimedia Systems, 2003, 9(4):353-364.
[3]. O.Boiman, M.Irani.Detecting irregularities in imagesand in video.Computer Vision, 2005, 1:462-469.
[4]. M. Cheng, G. Zhang, N.Mitra, X. Huang, S. Hu.Global contrast based salient region detection. IEEE Conference on Computer Vision and Pattern Recognition, 2011, 409-416.
[5]. D.Culibrk, M.Mirkovic, V.Zlokolica, M.Pokric, V.Crnojevic, D.Kukolj. Salient motion features for video quality assessment. IEEE Trans. Image Process, 2011, 20 (4): 948–958.
[6]. J. Findlay. The visual stimulus for saccadic eye movementin human observers. Perception, 1980, 9: 7-21.
[7]. D. Gao, S. Han, N.Vasconcelos. Discriminant saliency, the detection of suspicious coincidences, and applications to visual recognition. IEEE Trans. Pattern Anal. Mach. Intell, 2009, 31(6):989–1005.
[8]. J.Harel, C. Koch, P.Perona. Graph-based visual saliency. Advances in Neural Information Processing Systems, 2007, 19:545–552.
[9]. D.J. He, Y.M. Zhang, H.B. Song. A Novel Saliency Map Extraction Method Based on Improved Itti's Model. CCTAE, 2010, 323-327.
[10]. L.Itti, C. Koch, E.Niebur. A model of saliency-basedvisual attention for rapid scene analysis. IEEE Trans. Pattern Analysis and Machine Intelligence, 1998, 20(11):1254-1259.
[11]. L.Itti. Automatic foveation for video compression using aneurobiological model of visual attention. IEEETrans.Image Processing, 2004,13(10): 1304-1318.
[12]. S.Y.Kwak, B.Ko, H. Byun. Automatic salient-object extraction using the contrastmap and salient points. Advances in Multimedia InformationProcessing, 2004, 3332:138-145.
[13]. Z. Li, S. Qin, L.Itti. Visual attention guided bit allocation in video compression. Image Vision Compute, 2011, 29(1):1-14.
[14]. Y. Lin, B. Fang, Y. Tang. A computational model for saliency maps by using local entropy. AAAI Conference on Artificial Intelligence,2010, 967-973.
[15]. T. Liu, J. Sun, N. Zheng, X. Tang, H. Shum. Learning to detect a salient object. IEEE Conference on Computer Vision and Pattern Recognition, 2007, 1-8.
[16]. E.Rahtu, J.Kannala, M.Salo, J. Heikkila.Segmentingsalientobjectsfrom images andvideos. EuropeanConferenceonComputerVision, 2010, 366-379.
[17]. R.Valenti, N.Sebe, T.Gevers. Image saliency by isocentric curvedness and color. IEEE International Conference on Computer Vision, 2009, 2185-2192.
[18]. Y.Xue, Z. Liu, R. Shi. Saliency detection using multiple region-based features. Optical Engineering,2011, 50(5):057008-1-9.
[19]. A.A. Zaher, A.H. Ruba. Hill-manipulation:An effectivealgorithm for color image segmentation. Image and VisionComputing, 2006, 24:894-903.

Yongmei Zhang" Improved extraction approach of salient object based on visual attention mechanism"International Journal of Computational Engineering Research (IJCER), vol. 09, no. 1, 2019, pp 43-47