Salient object detection of fusing foreground seeds and center priori

Lingkang Gu*

School of Computer and Information, Anhui Polytechnic University, Anhui Wuhu China

*Corresponding author e-mail: glk_81@163.com

Abstract. In this paper, a salient object detection algorithm based on foreground seeds and center priori fusion is proposed. Firstly, using corner detection and edge linking algorithms obtains two convex hulls, and the approximate position of the target region is preliminarily determined by their intersection points. Then using the convex hull edge as the standard, the similarity of the hyperpixel in the convex hull is detected, and removing superpixels similar to most external edges. Finally, a center priori model fused with the foreground seeds, and the final significant image is obtained by using Bayesian optimization framework. By comparing and merging different visual features, it is shown that this proposed algorithm is very effective in the saliency object detection.

1. Introduction

Salience detection is one of the hotspots in computer vision [1]. It is widely used in moving detection and tracking, object extraction [2,3] and so on.

At present, salience detection algorithms can be broadly divided into two categories. One is top-down task-driven algorithm, the other is bottom-up data-driven algorithm. For complex scenes, especially when the basic attributes of the foreground target are grasped ahead of time, the top-down detection algorithm is usually used, which directly defines the foreground for significant detection [4], and the accuracy of detection is relatively high. Easier access to target characteristics; In this kind of algorithm, the key step is how to determine the target area more accurately.

Previous researches have been carried out, such as [4] using the harris algorithm to obtain convex hull to determine the target region, [5] expanding the convex hull range and searching for foreground seeds in convex hull. The corner detection of image and compressed image is carried out in reference to obtain two convex hulls, whose intersection is used to locate the target. Because only the gradient information of the image is considered in this method, the corner detection of the compressed image may make the location of the target region inaccurate when the compression cannot smooth out the gradient change.

In addition, most prominent targets are in or near the center of images. Therefore, central priori is often used in salient detection. Although most objects can be detected by a central priori model, there are some images in which the object is not in the center, which leads to the lack of an ideal detection result [6].
2. Our algorithm

In order to solve the above problems, based on the previous studies, we propose a salient object detection method based on the fusion of foreground seeds and center priori. It consists of three parts:

1) Firstly, the harris corner of the input image [4] is found, and the convex hull is constructed to obtain the foreground seed.

2) Secondly, a central priori model was constructed and fused with foreground seeds to form a priori image.

3) Using Bayesian formula to construct the final salient image:
   A) The priori probability of Bayesian formula is taken as the salient value of the priori image;
   B) Using the constructed convex hull to calculate the likelihood probability of Bayesian formula;
   C) Finally, the priori probability and the likelihood probability are put into Bayesian formula to calculate, and the calculated probability is taken as the salient value of the pixel to obtain the final significant image.

3. Foreground seed

In a visual scene, the most attractive area is the most representative region, which is the significant image element in salience detection, which is also called foreground seed. On the basis of foreground seeds, they spread to the surrounding area until the whole significant region. Therefore, it is very important to locate the seeds accurately. There are many methods to select foreground seeds, the most common and simple method is convex hull [7-9].

3.1. Calculation of convex hull

In this paper, harris operator [4] is used to estimate corner points and contour points, and edge interference points are removed by threshold method to obtain convex hull. However, using corner detection is not enough for salient detection, and the location information of object should be considered. Therefore, in order to locate the object region more accurately, the convex hull of the image is calculated from two different angles of corner detection and position information.

The main part of the object is inside the convex hull, while the outer part is mainly the background area. However, there are still many backgrounds in the convex hull. In order to solve this problem, the compressed image of the original image is used in the paper to detect the secondary corner, and the new convex hull is obtained to improve the original convex hull. The convex hull obtained is closer to the target. However, for the object on the edge, it will not be smoothed out because of the compression, so the influence of the diagonal point still exists, making the convergence of convex hull to the target is not ideal in this case.

For the above reasons, in order to obtain the minimum convex hull which is closer to the target, the concept of edge join is introduced in this paper, and the position feature is added to calculate the convex hull. Definitions of edge connections as Eq. (1):

\[
\text{bdC}_i = \frac{\text{len}(R_i)}{\sqrt{\text{Area}(R_i)}}
\]

In which, \(\text{len}(R_i)\) represents the number of pixels on the image edge of the \(R_i\) in the \(i\) region, \(\sqrt{\text{Area}(R_i)}\) is the area of the region \(R_i\). \(\text{bdC}_i\) represents the edge join value of the \(i\) region \(R_i\).

3.2. Selection of foreground seeds

Based on the minimum convex hull \(c\), the foreground region \(Fc\), is estimated. Rough foreground seeds are obtained by collecting superpixels satisfying the conditions in \(Fc\). In order to further determine the foreground seed, the adjacent super-pixels are added to \(c\), and the background region which is closer to the target region around the outer edge of \(c\) is defined, and a foreground seed is defined as:

\[
\text{Sim}_i = \frac{1}{\sum_{k \in \{L, R, B\}} (c_i^k - c_j^k)^2}
\]
in which, \( c_i \) is the value of the super pixel \( i \) on the \( k \) th color channel. When the superpixel \( i \) in \( c \) is not similar to most of the superpixels in \( c' \), it can be identified as a foreground seed.

4. Center prior model

Because the image object is usually located in the center or near the center of the image, a priori model of the initial center is constructed, which lightens the image region which the convex hull fails to encircle. But it also lights up non-significant areas, so the results are still unsatisfactory. For this reason, color features are added to the initial central priori model to construct a central priori model, which not only lightens the prominent region of the image, but also inhibits the non-significant region, which makes the detection results more accurate.

Constructing a preliminary central priori model:

\[
C(i) = \exp\left(-\frac{(x_i - x_0)^2}{2\sigma_x^2} - \frac{(y_i - y_0)^2}{2\sigma_y^2}\right)
\]  

(3)

in which, \( x_i \) and \( y_i \) represent the average horizontal value and the average vertical coordinate value of the super-pixel \( i \), and \( x_0 \) and \( y_0 \) represent the horizontal coordinate value and the vertical coordinate value of the center point of the input image. Parameters \( \sigma_x \) and \( \sigma_y \) mean horizontal and vertical variances, respectively, in this paper sets \( \sigma_x = \sigma_y = 0.5 \) and normalizes the pixel coordinate value to \([0,1]\).

For each super-pixel in the image, an \( n \times n \) correlation matrix is constructed by using the cieL*a*b* color feature:

\[
R(i,j) = t(i,j)
\]  

(4)

Where \( i,j \in [1,N] \), \( N \) is the number of superpixels in the image; \( t(i,j) \) denotes the Euclidean distance between the color of the first and the \( j \) super pixels, and sets \( t(i,j) = 0 \). Calculate the sum of each column in \( R \) and get a vector:

\[
V = \{ \sum_{i=1}^{N} t(i,1), \sum_{i=1}^{N} t(i,2), \ldots, \sum_{i=1}^{N} t(i,N) \}
\]  

(5)

The vector \( v \) is normalized and the value in \( v \) is used as the weight of each superpixel. The central priori model of this paper is constructed:

\[
C_i = v_i \times c(i)
\]  

(6)

Where \( c(i) \) is the primary central priori model, and \( v_i \) denotes the first term in the vector \( v \).

The foreground seed was fused with the central priori model, and a priori image was obtained [6]. In the prior image, \( s_i \) as the pixel salient value:

\[
S_i = Sim_i \times (1 - e^{-\beta C_i})
\]  

(7)

In which, the seed of foreground is \( Sim_i \), \( C_i \) as the central priori model, and \( \beta \) as an equilibrium parameter.

5. Likelihood probability model

Usually, the significant value of pixels in convex hull is relatively high, but in some cases, the color of parts of prominent targets in the image is not obvious, so it is easy to be used as background area in detection. However, pixels in some background which are similar to foreground color and have significant color may be mistaken for salient object by the algorithm. Therefore, the likelihood probability inside the convex hull is considered as the likelihood probability of the salient object \( p(v|l) \) and the likelihood probability outside the convex hull as the likelihood probability of the background \( p(v|b) \).
For the number of pixels, the pixel \( v \) of the image, \( n_{out} \) is the number of pixels outside the convex hull, and \( N_{in}(f(v)) \) and \( N_{out}(f(v)) \) are the statistical values of the color inside and outside the convex hull, respectively. Similar to reference [5], the likelihood probability of the whole graph is calculated:

\[
p(v|l) = \prod_{f \in [L,a,b]} \frac{N_{in}(f(v))}{n}
\]

\[
p(v|b) = \prod_{f \in [L,a,b]} \frac{N_{out}(f(v))}{N_{out}}
\]

Where \( L, a, b \) is an independent feature in the CIELab color model.

For the priori image \( s_i \), obtained by Eq.(7), the value of \( s_i \) is regarded as the priori probability of the first super-pixel \( p(l) \), for each pixel of the image, the priori probability is equal to the priori probability of the super-pixel in which the pixel is located. Then the priori probability of the pixel and the likelihood probability of the pixel in the image are put into the Bayesian formula and the calculated probability is taken as the significant value of each pixel so as to obtain the final significance diagram. Bayesian Model Formula used in this algorithm:

\[
p(l|v) = \frac{p(l)p(v|l)}{p(l)p(v|l) + p(b)p(v|b)}
\]

![Comparison of detection results between different algorithms on two datasets.](image)

**Figure 1.** Comparison of detection results between different algorithms on two datasets.

The Eq.(8) and Eq. (9) is put into the Eq.(10) to obtain the significant value of each pixel in the graph. Where, \( p(*) \) denotes probability, \( p(l) \) denotes a priori probability of significance.

6. Simulation experiment and result analysis

The experimental environment of this paper is CPU intel core 2.90 GHz, memory 16GB, SSD hard disk, and the experimental platform is Matlab, test dataset ECSSD and MSRA 10k [7]. MSRA 10k contains a relatively simple background, and ECSSD contains rich colors, diverse textures and complex backgrounds. The six significant detection algorithms are: HC [7], CTC [8], SCUL [9], HCCH [10], MDBL [11], ODSR [12].

Figure 1 contains a sample of salient images generated on MSRA 10k and ECSSD by this algorithm and six other algorithms. As can be seen from Figure 1a, because of the simple background of MSRA 10k, the performance of this algorithm is better than other algorithms. It can be seen from Figure 1b that on the ECSSD dataset, the performance of this algorithm is also the best, and the significant object in the image can be detected well.

6.1. Result analysis

There are many indexes to evaluate the effectiveness of salience detection algorithm. The commonly used indexes are \( PR \), \( F\text{-measure} \) and \( \text{Mean Absolute Error} \) (MAE) [13].
6.1.1 Precision-Recall (PR) curve. The PR curve is the most commonly used method to evaluate the effectiveness of the algorithm. The $P$ and $R$ are shown in Eq. (11) and Eq. (12), respectively.

$$
P = \frac{|B \cap G|}{|B|} \quad (11)
$$

$$
R = \frac{|B \cap G|}{|G|} \quad (12)
$$

Among them, $G$ is the significant truth image of the artificial marker; $B$ is the binary mask image generated by the threshold saliency image $S$; and the image area is represented as $|\cdot|$. If than the segmentation threshold, the pixel significant value is greater, the pixel is marked as a significant foreground, otherwise marked as a background. While the threshold is changed from 0 to 255, a complete PR curve can be drawn to evaluate the performance of the algorithm in different states, as shown in figure 2a, 2b. It can be seen from these two images that our algorithm is comparable to two excellent algorithms (ODSR and MDBL). Our experimental results show that on MSRA 10k the proposed algorithm has the best performance and on ECSSD is comparable to ODSR.

6.1.2 F-measure ($F_\beta$). In this paper, the Precision and Recall are weighted and averaged by F-measure value, as shown in Eq. (13):

$$
F_\beta = \frac{(1+\beta^2) \times P \times R}{\beta^2 \times P + R} \quad (13)
$$

in which $\beta^2=0.3$, the higher the $\beta$, the better the effect of the algorithm. We can see from the following Figure 3, that on MSRA 10k the algorithm has the best $F_\beta$ value, the ODSR is inferior to it, and the better $F_\beta$ value on ECSSD is similar to MDBL, as shown in Figure 3. These results further prove the validity of our proposed method.
6.1.3 Mean absolute error (MAE). In order to describe more accurately the consistency between salient images and truth images, MAE is introduced as an evaluation supplementary metric to calculate the errors between salient images $S$ and marked significant truth images $G$, as shown in Eq.(14):

$$\text{MAE} = \frac{1}{W \times H} \sum_{i=1}^{W} \sum_{j=1}^{H} |S(i,j) - G(i,j)|$$  \hspace{1cm} (14)

In which, $S(i, j)$ is the significant value, $W$ represents width and $H$ represents height of the image. The experimental results show that the MAE of the proposed algorithm is the smallest in two databases, which indicates that the proposed algorithm is superior to other algorithms and has a high accuracy.

| Algorithms | MSRA 10k | ECSSD | Time/s |
|------------|----------|-------|--------|
| HC         | 0.317    | 0.297 | 0.203  |
| CTC        | 0.281    | 0.274 | 0.317  |
| Our        | 0.109    | 0.112 | 0.218  |
| SCUL       | 0.218    | 0.211 | 0.241  |
| HCCH       | 0.176    | 0.192 | 0.352  |
| MDBL       | 0.107    | 0.113 | 0.479  |
| ODSR       | 0.102    | 0.114 | 0.375  |

6.2. Algorithm efficiency

It can be seen from Table 1 that the running time of this algorithm is moderate. Particularly for the better MDBL algorithm and the newer ODSR algorithm, the speed of this algorithm is much faster than them. For faster algorithms (such as HC), the effect of this algorithm is far ahead. The speed of the proposed algorithm is faster than that of MDBL and ODSR algorithm which run well. But for faster algorithms (such as HC), the effect is better.

7. Conclusion

In this paper, a salient object detection algorithm is proposed. Firstly, we construct the seeds of foreground. Secondly we combine the central priori and Bayesian optimization framework to obtain the final significant image. The experimental results show that the proposed algorithm highlights the significant object, and embodies the superiority of the algorithm. By comparing and merging different visual features, it is shown that this proposed algorithm is very effective in the saliency object detection.

Acknowledgements

This work was supported by “Undergraduate Teaching Quality Improvement Program of Anhui Polytechnic University (No. 2018jyxm45)“.

References

[1] Wang Gang, Zhang Yongdong, Li Jintao. High-level back ground prior based salient object detection[J]. Journal of Visual Communication and Image Representation, October 2017, Vol.48, pp.432-441.

[2] Qin C C, Zhang G P, Zhou Y C, et al. Integration of the saliency based seed extraction and random walks for image segmentation[J]. Neurocomputing, 2014, 129: 378-391.
[3] Wang, Xiang ; Ma, Huimin ; Chen, Xiaozhi ; You, Shaodi. Edge Preserving and Multi-Scale Contextual Neural Network for Salient Object Detection[J]. IEEE Transactions on Image Processing, Jan. 2018, Vol.27(1), pp.121-134.

[4] XIE Yulin, LU Huchuan, YANG Minghsuan. Bayesian saliency via low and mid level cues[J]. IEEE Transactions on Image Processing, 2013, 22(5): 1689-1698.

[5] LIU Risheng, CAO Junjie, LIN Zhouchen, et al. Adaptive differential equation learning for visual saliency detection[C]. Computer Vision and Pattern Recognition 2014, Columbus, USA, 2014: 3862-3869.

[6] Huo L, Jiao L C, Wang S, et al. Saliency detection with color attributes [J]. Pattern recognition, 2016, 49:162-173.

[7] Cheng M M, Mitra N J, Huang X L, et al. Global contrast based salient region detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 569-582.

[8] Qing Zhang, Jiajun Lin, Yanyun Tao, Wenju Li, Yanjiao Shi. Salient object detection via color and texture cues[J]. Neurocomputing, 2017, Vol.243, pp.35-48.

[9] Zhang Qiang, Liu Yi, Zhu Siyang, Han Jungong. Salient object detection based on super-pixel clustering and unified low-rank representation[J]. Computer Vision and Image Understanding, 2017, Vol.161, pp.51-64.

[10] Liu Qing, Hong Xiaopeng, Zou, Beiji, et al. Hierarchical Contour Closure-Based Holistic Salient Object Detection[J]. IEEE Transactions on Image Processing, 2017, Vol.26(9), pp.4537-4552.

[11] Song Hangke, Liu Zhi, Du Huan, et al. Depth-Aware Salient Object Detection and Segmentation via Multiscale Discriminative Saliency Fusion and Bootstrap Learning[J]. IEEE Transactions on Image Processing, 2017, Vol.26(9), pp.4204-4216.

[12] Yan Xiaoyun, Wang Yuehuan, Song Qiong, et al. Salient object detection via boosting object-level distinctiveness and saliency refinement[J]. Journal of Visual Communication and Image Representation, 2017, Vol.48, pp.224-237.

[13] Georgia Gkioxari, Ross Girshick, Piotr Dollár, and Kaiming He. Detecting and Recognizing Human-Object Interactions. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.