IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 25, 2018, accepted January 7, 2019, date of publication January 11, 2019, date of current version February 8, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2892558

Efficient Visual Saliency Detection via Multi-Cues

JUNFENG WU[1,2,3], HONG YU[2], JIANWEI SUN[4], WENYU QU[1], AND ZHEN CUI[2]
[1]School of Computer Science and Technology, Tianjin University, Tianjin 300072, China
[2]College of Information Engineering, Dalian Ocean University, Dalian 116023, China
[3]Guangdong Province Key Laboratory of Popular High Performance Computers, Shenzhen 518000, China
[4]Beijing Baidu Netcom Science and Technology Co., Ltd., Beijing 100089, China

Corresponding author: Wenyu Qu (wenyu.qu@tju.edu.cn)

This work was supported in part by the Guangdong Province Key Laboratory of Popular High Performance Computers under Grant SZU-GDPHPCL201805, in part by the Institute of Marine Industry Technology of Universities in Liaoning Province under Grant 2018-CY-34, in part by the National Natural Science Foundation of China under Grant 61701070 and Grant 61672379, in part by the Liaoning Doctoral Start-up Fund under Grant 20180540090, in part by the National Natural Science Foundation of China-Guangdong Joint Fund under Grant U1701263, and in part by the China Postdoctoral Science Foundation under Grant 2018M640239.

ABSTRACT Saliency detection has become a valuable tool in computer vision processing, which has been attracting a good deal of attention. Although a lot of research has been done, it cannot obtain ideal performance. In most saliency detection framework, the single difference cue is often used to detect saliency map, but their results were far from satisfaction for low discriminative power of single cue. In addition to the composition of the detection cues during the whole process, the used cues which could only describe the low level information could also lead to poor detection. In order to solve this problem effectively, the paper proposed a novel saliency detection framework by fusing multi-cue difference with high level difference. Basically, the comprehensive information is coupled into multi-cues vectors to remove the non-salient regions and enhance the brightness of the salient area. Specifically, we utilize the fusion of multi-cues and high-level information to improve the ability of understanding images. To further improve the performance of the proposed framework in AUC, we adopt multiple assignments strategy while enhancing the precision of saliency detection. Extensive experiments indicate that the newly constructed multi-cues with high level information could effectively suppress the influence of background information on salient regions. In order to verify the effectiveness of the new algorithm, our experiment uses several standard benchmark datasets (MSRA, ASD, SED1, SED2, and SOD) to test the performance of the algorithm. The experimental results demonstrate that this method has achieved good saliency detection result and good AUC performance in the test. More importantly, the experimental results also show that the proposed method is well complementary to many existing algorithms.

INDEX TERMS Saliency detection, pattern distance, high-level, multi-cues.

I. INTRODUCTION

With the development of computer vision application, saliency detection is attracting more and more attention. Saliency detection is to detect the salient regions or pixels in the images, which has become a valuable tool in many computer vision or artificial intelligence tasks due to its broad application, such as object detection [1]–[3], search engines [4], [5], action recognition [6], [7] and object tracking. Image segmentation is similar to saliency detection, but saliency detection is more practical and operable compared with image segmentation [8], [9]. Accurate image segmentation and edge extraction in practical systems are still open problems. It is obvious that the saliency detection results should conform to the visual characteristics of the human eye [10], this also means that the results of detection must be consistent with those observed by human eyes.

For some time, more and more experts and scholars have been devoting themselves to this domain, and they have made many great achievements. Generally speaking, there are two emerging trends for the cues used for saliency detection, one is to use the low level cue in the images to detect saliency; the other would be the use of high level information. The low level cue of an image mainly refers to the information at the pixel level [11], most saliency detection with low level cues mainly focuses on the change of pixels and the relationship between pixels. Compared with low level cues, the high level cues pay more attention to a wider scope in the images; semantic information and relationship between regions in images is also taken into account. For most algorithms with high level cues, the region segmentation results are vital to the saliency detection. A lot of experiments have proved that these measures improve the detection results, but the final experiment results are affected seriously by region segmentation under the situation of high level cues [12]. Therefore, many scholars are dedicated to the research of region segmentation in order to improve the performance of the saliency detection. In many methods, super pixel algorithm performs very well [13]–[15], which has been used in many saliency detection methods. However, we find that the irrelative information in images will cause serious interference to the detection, which could reduce the detection quality, and make some processing unstable or failed. Furthermore, background information in the images can interfere with the detection of salient regions too. After a long-term research in depth, we find that two problems should be solved if we want to get a better detection performance. First of all, we should take appropriate measures to minimize the interference of background information to saliency detection. Only the interference is eliminated, the saliency area will be highlighted. Secondly, the distinction between the salient region and background information should be increased. With the increase of discrimination, the results of saliency region extraction will be more accurate.

In this paper, we have taken into full consideration of the situation above, we gives a new saliency detection method by high level multi cues (HMDS). The paper reduces the luminance of background by the using PCA space and Gaussian prior weight. In order to further distinguish salient regions from their backgrounds, object level regions are utilized in the process of saliency detection.

To verify the effectiveness of the proposed algorithm, we tested the proposed method on some well-known data sets, including MSRA, ASD, SED1, SED2 and SOD. The simulation experiment results show that the proposed method performs correct and effectiveness.

The rest of this paper is organized as follow. We discuss the related work in this field in section 2. The HMDS algorithm is described in section 3.In section 4, the simulation results will be shown in detail. Finally, we summarize our paper and the future work is also prospected.

II. RELATED WORK

With the development of saliency detection, more and more attention has been paid to it. Many scholars have proposed many algorithms to improve the performance of saliency detection. Achanta [16] presented a saliency detection method with pixel level. This algorithm is effective and takes into account the global differences. The algorithm also proposed a new evaluation method and got good performance. Cheng [8] utilized color histogram contrast to detect salient objects in the images, which was still at the pixel level. The simulation results show that the method has high efficiency and can have good performance in MSRA dataset. Unlike pixel level methods, more and more methods attempt to detect saliency map at a higher level, which keep their interest focused mostly in the regions and objects in the images [8], [14]. Goferman *et al.* [17] proposed two saliency map, which are pixel level saliency map and region level saliency map respectively. The paper fused two saliency maps together and proposed a new algorithm to further optimize the final result. The experimental results indicate that combination method dose seem to outperform the previous algorithms, but background removal is not complete and the foreground are not evenly distributed. Jiang *et al.* [12] attempted to detect several significant regions with super pixel segmentation, which is called CBS. This method could distinguish salient region from image obviously and the gray value of salient region is more uniform. But in the final result, the residue of background is relatively obvious, the background in the images are not completely eliminated. Ran *et al.* [18] argued the image blocks distributed in different locations can influence the performance of saliency detection. So they proposed a saliency detection method by calculating color and pattern difference. In this paper, the background in the images is thoroughly cleared but the luminance of foreground is reduced at the same time.

Through the analysis of previous work, we find that both pixel level and object level could influence the performance of saliency detection. Through further research, we also found that the image blocks used in saliency detection can also directly affect the results of saliency detection. For full consideration mentioned above, this paper proposed a new saliency detection method with high level multi cues (HMDS).

III. THE PROPOSED METHOD

As mentioned above, compared with low level differences, high level differences are more abstract and difficult to extract, but high level differences contain a lot of useful information that can be used to detect saliency. Therefore, it will be very helpful for saliency detection if the high level differences can be calculated accurately. Generally speaking, the information contained in an image is closely related to its entropy. There is substantial research indicating that the effective information contained in a thing increases with the decrease of entropy. This also means that high level differences can be better used to detect the saliency map. So in this paper, we attempt to utilize high level difference to detect saliency map rather than low level difference. So compared with low level differences, we prefer to use high level differences to detect saliency. Pixels and relationships between pixels are often used to detect low level saliency. Unlike pixels, the image patches are local areas around pixels, which contain a lot of information, including location information, semantic information, and gray value change information and so on. Obviously, compared with the pixel method, the saliency map detected by image block could be more meaningful. Because image blocks may contain certain semantic information, so image blocks can describe the whole image more objective. Therefore, the saliency detection based on image blocks should be better than that of pixels. If we do further research, we find that the information contained in the image block is incomplete too. After all, the image patches are in a local area. The information that can be described is very limited. Object-level differences contain more semantic information than block-level differences. So it would be more reasonable to detect saliency by using object level difference.

For low level difference, it is not so easy to calculate object level difference because extracting meaningful regions from images is not an easy task [12]. This poses a great challenge for us to utilize object level differences to detect saliency map. From another perspective, the background information of the image will seriously affect the extraction of objects in the image. A lot of experiments show that pixels belonging to the same object may also have different saliency score in different background. In our work, we try our best to eliminate the influence of background on useful objects. The image pattern is integrated into the whole saliency detection framework and used to reduce the influence of background [18]. The closed boundary in the images is also the focus of our attention, which may contain objects. In order to further optimize the detection results, we introduce high level cues in to our work. Over these considerations, we propose a novel saliency detection method to pursue a better saliency detection performance.

A. THE DIFFERENCE OF PATTERN DISTANCE AND COLOR DISTANCE

Through profound and careful observation and analysis, we found that salient patches and non-salient blocks differ greatly in spatial distribution. Most of the non-salient patches are clustered in the high-dimensional space of the image, while the salient patches scatter throughout the image at all locations. During the process of saliency detection, we need to use average patches which are deeply influenced by patches. The pattern distance is the distance between the patches and the principal direction, which was calculated firstly. The calculation of pattern distance is carried out in PCA coordinate system. The pattern distance under PCA could help us remove the background in the images effectively.

As for color difference, in order to conform to the visual characteristics of the human eye, the paper utilizes LAB color space to calculate color distance between patches (Fig.1.c).

The distance is defined as follow [18]:

$$PD\left(p_x\right) = P\left(p_x\right) \cdot C(p_x) \tag{1}$$

where p_x is the patch in the images, $P(p_x)$ is the pattern distance of image patches, $C(p_x)$ is the color distance of image patches. $PD(p_x)$ is saliency map for each image patch.

Generally speaking, people like to put objects in the center of the image when take pictures. So the patches near the center are more important than the patches far from the center. Based on such considerations, the paper further uses the Gaussian weighting formula to adjust the weight of each image patch during the whole process. For Gaussian weighting formula, the closer to the center, the greater the weight, so we can increase the weight of the pixels close to the center in this way.

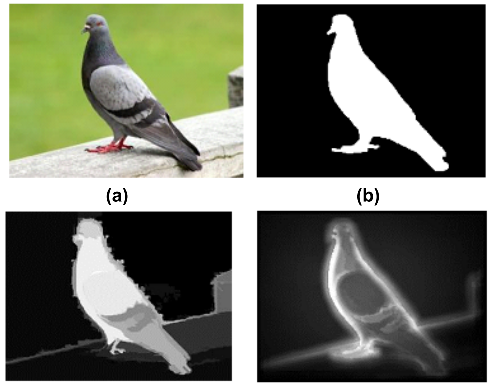


FIGURE 1. Saliency detection results of different methods. (a) Original image. (b) Ground truth. (c) Pattern distance result. (d) Object level result.

$$PS\left(p_x\right) = PD\left(p_x\right) \cdot G(p_x) \tag{2}$$

where $G(P_x)$ is Gaussian weighting formula, which is generated by $PD(p_x)$. And $PS(p_x)$ is saliency score obtained by weighted calculation

B. THE DIFFERENCE OF OBJECT LEVEL

After the analysis above, we know that it is possible to include meaningful objects in closed curves in images. So if we can find the closed curve in the image, it is significant for saliency detection. One of the effective ways to find the closed area in an image is to use image segmentation algorithm. The super pixel method is classical and effective, which is widely used in image retrieval, object recognition and anomaly detection. In our work, super pixel algorithm is also used to obtain closed regions in images through multiple iteration. After the closed region is obtained, the next step is to calculate the distance between regions; we could calculate the distance between the regions as follows:

$$OD\left(r_i\right) = \sum_{j=1}^{k} \alpha d_{color}\left(r_i, r_j\right)(1 \le i \le n) \tag{3}$$

where r_i and r_j is region detected by super pixel algorithm, which may contain meaningful objects. r_j is one neighbor of r_i, d_{color} stand for the color difference between the detected regions. α is a constant.

And then the saliency map $OS(r_i)$ with closed region could be got as (4)

$$OS\left(r_i\right) = -w_i ln(1 - OD(r_i)) \tag{4}$$

where w_i is the weight arranged for each region. Through experiments we find that the saliency score of the pixels in the region increases significantly with the increase of color difference.

C. THE SALIENCY DETECTION VIA HIGH LEVEL MULTI-CUES

By means of above process, we could get two differences, which are pattern and color difference and object level difference. Each difference has own advantages and characteristics, which could play different roles in saliency detection process.

With the pattern and color difference, we can get ideal saliency map, the results are accurate and adaptable. But the prominent problem with this difference is that the distinction between background and background is not obvious. Through analysis, we find the main reason is that the uses of PCA coordinate system. In the process of calculating the difference with PCA coordinate system, a lot of information is loss compressed. This leads to a lot of information we get is not complete.

With object level difference, we can get saliency maps with distinct distinctions between foreground and background. However, the brightness of the background has not been completely suppressed and the brightness of the foreground has not been improved. On the other hand, although more semantic information can be obtained by detecting saliency maps with regions, but the influence of the background on the foreground has also been completely ignored. From the perspective of saliency detection, this is also unreasonable.

In order to solve these problems, we have taken a series of measures. Through the complementing each other of these two kinds of differences, the two optimization differences could strengthen each other thus forms a satisfactory algorithm with excellent performance. Another measure is to give different weights to different differences according to different needs.

The first thing we need to do is to suppress the background and highlight the brightness and discrimination of the foreground. We normalize the pattern and color difference and object level difference:

$$PS^* = PS/max(PS) \tag{5}$$

$$OS^* = OS/max(OS) \tag{6}$$

After normalizing the two differences respectively, two differences are put together for saliency detection:

$$HS_{tmp1} = PS^* \cdot OS^* \tag{7}$$

Through many experiments we find that putting the two differences together indiscriminately is unreasonable. It's hard to make full use of the advantages of their differences. So we should make rational use of the two differences according to different emphases:

$$HS^{*i}_{tmp2} = \begin{cases} w_1 \cdot HS^{*i}_{tmp1} & if(HS^{*i}_{tmp1} \geq m) \\ HS^{*i}_{tmp1} & if(HS^{*i}_{tmp1} < m) \end{cases} \tag{8}$$

where HS^*_{tmp2} stands for the saliency score we obtained from the previous step, m stands for the average value of HS^*_{tmp1}. w_1 is the weight of HS^*_{tmp1} and it is utilized to adjust the value of saliency score. In this paper, the value of w_1 is set to *2* through repeated experiments

FIGURE 2. Saliency detection results: the first row is original images, the second row is ground truth and the third row is saliency map.

TABLE 1. Configuration of experimental platform.

Category	Configuration
CPU	Intel(R) Core i3-4150 3.5GHZ
main memory	4.0GB
hard disk	500GB
OS	Windows 7
environment	MATLAB 2012B

In order to further improve the brightness of the foreground and suppress the background information in the final result, the color and pattern difference is utilized again. The main reason is that color difference can significantly improve the brightness of foreground information.

$$HS = w_2 \cdot HS^{*i}_{tmp2} + w_3 \cdot PS^* \tag{9}$$

In which w_2 and w_3 are the weight of HS^{*i}_{tmp2} and PS^* respectively. In this paper, the value of w_2 is set to *1* and the value of w_3 is set to *0.5*. These values are empirical values obtained through repeated experiments. We can set different values for these parameters according to different situations. Generally speaking, the values of w_1 and w_2 are greater than *1*, and the value of w_3 is less than *1*. BecausePS^* is used to suppress the brightness of the background and the ultimate saliency score should not be too great, so w_2 and w_3 are used to adjust the relationship betweenHS^{*i}_{tmp2} and PS^*. The purpose of setting these weights is to further improve the performance of the algorithm; another purpose is to prevent the saliency score growing greater.

In conclusion, we can get the final saliency map through the above operations (Fig.2).By observing the experimental results we find that the saliency maps are very similar to those of ground truth. Foreground information in saliency maps has sufficient brightness; the background is also well suppressed.

IV. EXPERIMENT

To test the effectiveness of our proposed algorithm, we do a series of experiments on some famous datasets to test our algorithm (HMDS). The software and hardware settings of the experiment platform are as follows.

Dataset: In order to evaluate our proposed algorithm objectively, we test our method on five popular datasets.

Each dataset consists of the original images and ground truth images. The original image is used for testing; the ground truth images are used to compare with the results. The details of each data set are as follows.

MSRA dataset [19]: There are 5,000 images in this dataset, including the original images and ground truth images. This dataset is manually labeled by nine people. It is one of the most popular datasets in the field of computer vision.

ASD dataset [16]: There are 5,000 images in this dataset. It consists of 1,000 original images and ground truth images selected from MSRA datasets.

SED1 dataset [20]: There are 100 images in this dataset. There is only one object in each image.

SED2 dataset [20]: There are 100 images in this dataset. There are two objects in each image.

SOD dataset [21]: There are 300 images in this dataset. In order to eliminate human subjectivity as much as possible, the ground truth images of this dataset are completed by seven different people.

Baseline methods: To evaluate our algorithm objectively, seven popular saliency detection algorithms are selected and compared with our algorithms: PCAS [18], CBS [12], FT [16], HC [8], LC [11], RC [8] and SR [16]. The codes of these algorithms are downloaded from open links provided by the authors. We use the same software and hardware configuration when testing these algorithms.

Evaluation system: In order to objectively and comprehensively evaluate the proposed algorithm, we designed a series of experiments. The precision-recall curve is one of the important indicators that can reflect the performance of the algorithm. The threshold value used for segmenting the saliency map is from 0 to 255. In order to test the accuracy of the algorithm, the experiment results are compared with ground truth on five different datasets. As we know, higher precision and recall rate are important indicators of a good algorithm, so F-measure curves is used as final evaluation index [1]. In order to make the evaluation system of algorithm more comprehensive, we introduce AUC (Area-Under-The-Curve) to evaluate the performance of our algorithm on five selected datasets. In order to make the performance of the algorithm more intuitive and understandable, Mean AUC is used to describe the average robustness and precision of the proposed algorithm on the whole testing dataset. Finally, we show the performance of the proposed algorithm in terms of operation running time.

HMDS validation: We show some comparisons the results with seven other algorithms on five datasets (Fig.3). By observing the experiment results, we can find that the brightness of foreground is effectively enhanced. Compared with other algorithms, the background is effectively suppressed and the brightness is lower. The distinction between foreground and background is more obvious and the edge is clearer. Generally speaking, the proposed algorithm is superior to the other seven algorithms.

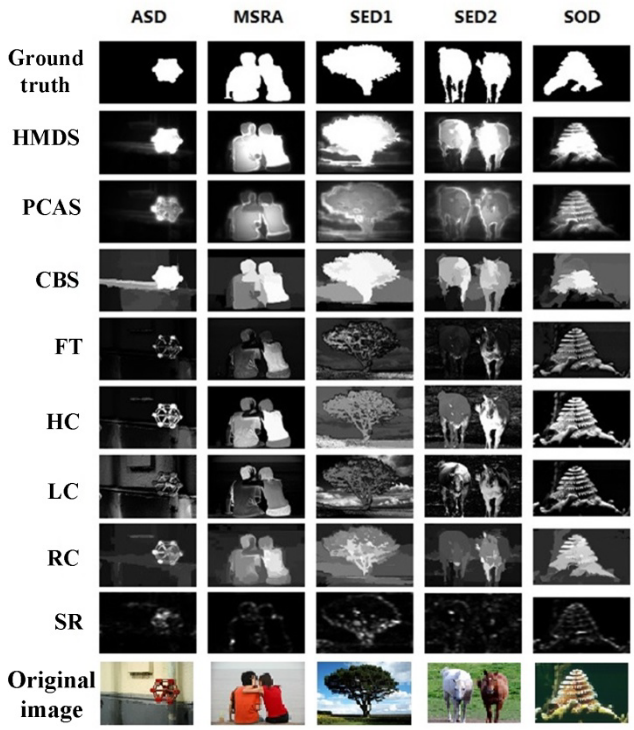


FIGURE 3. Saliency map of different methods.

Fig.4 and Fig.5 show the performance of seven different algorithms on precision and F-measure in five datasets. We can find that our algorithm has excellent performance in precision and F-measure. And in a wider range, the F-measure of HMDS algorithm is more ideal on the selected datasets than other seven methods. In particular, our algorithm performs better than the other seven methods if the threshold is less than 0.5. This also indicates that our algorithm has an extensive applicability and could be used in different applications.

In order to evaluate the effectiveness of our proposed algorithm objectively and comprehensively, we list the AUC (Fig.6) and mean AUC (Fig.7) about eight methods in 5 different datasets. From Fig.5 we can see that the proposed method has the greatest AUC value in each dataset. Mean AUC is the sum of AUC values obtained from five data sets, and then the average value is calculated. Mean AUC can reflect the comprehensive performance of the algorithm in different datasets, which can better reflect the robustness of a saliency detection algorithm. From Fig.6 we can know that compared with the other seven algorithms, our algorithm performs better in mean AUC, which means that it is more robust.

In conclusion, a large number of experimental results the proposed algorithm can better highlight the foreground information in the image; it can also suppress the influence of background on the foreground in the detection process. Moreover, it is more robust and adaptable on different datasets.

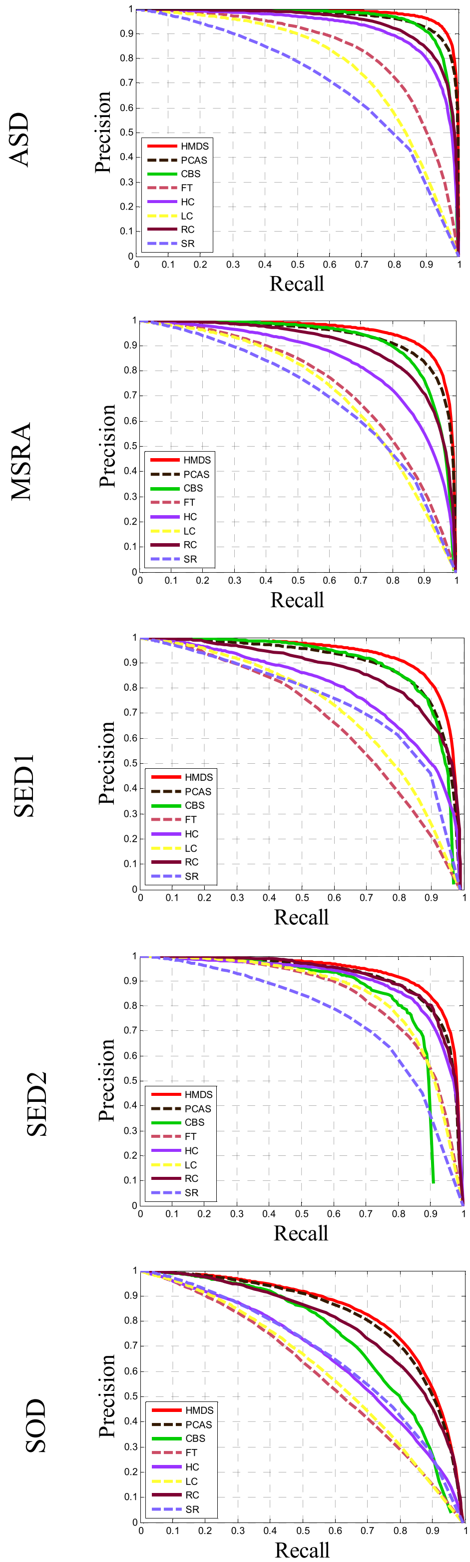


FIGURE 4. The precision comparison on different datasets.

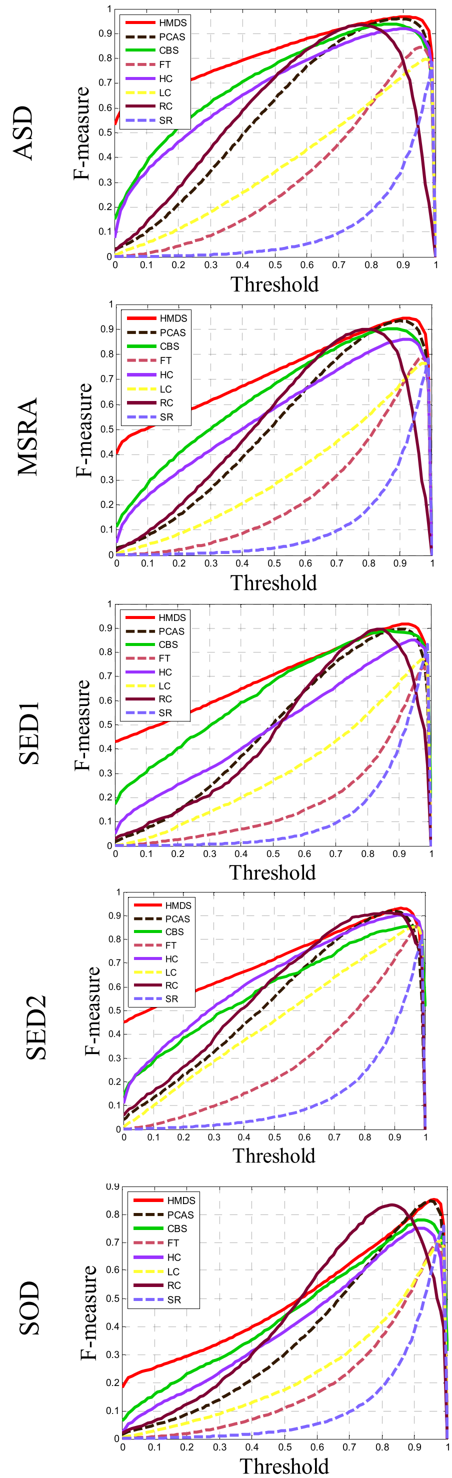


FIGURE 5. The F-measure comparison on different datasets.

Running time: The main purpose of this algorithm is to solve the problem of precision and robustness in saliency detection. The running time increases with the complexity of the algorithm. When testing the precision and robustness of the algorithm, we also give a comparison between the proposed algorithm and other algorithms in time complexity. All methods are implemented in the same software and hardware environments. The time required to detect images in different datasets is different. Compared with other algorithms, our algorithm does not perform well in computing time. This is because our algorithm is more complex than

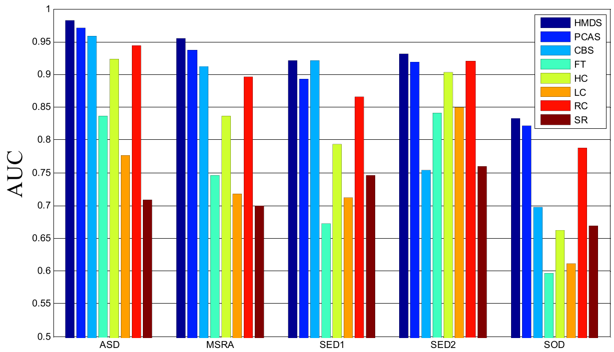


FIGURE 6. The AUC result.

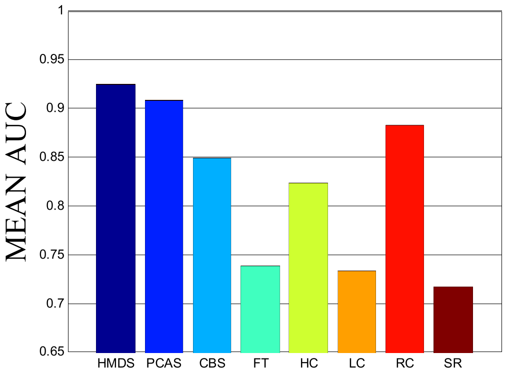


FIGURE 7. The mean AUC result.

other algorithms. It takes about 3.961 seconds to compute an image saliency map.

V. CONCLUSION

In order to solve the problem of low robustness and accuracy in saliency detection algorithm, the paper proposed a novel saliency detection framework. The algorithm uses multi-cues to improve the accuracy and robustness of the algorithm. The paper also adopts multiple assignments strategy to enhance the precision of saliency detection. Extensive experiments indicate that the newly constructed multi-cues with high level information could effectively suppress the influence of background information on salient regions.

In order to verify the effectiveness of the new algorithm, our experiment uses several standard benchmark datasets (M S R A, ASD, SED1, SED2 and SOD) to test the performance of the algorithm. The experimental results demonstrate that this method has achieved good saliency detection result and good AUC performance in the test.

In the future work, we will focus on using neural networks to solve the problems in saliency detection. We will also further strengthen the research on saliency detection of low contrast images and underwater images. And we intend to apply saliency detection in image retrieval, underwater target detection and other fields.

REFERENCES

[1] K. Ishikura, N. Kurita, D. M. Chandler, and G. Ohashi, "Saliency detection based on multiscale extrema of local perceptual color differences," *IEEE Trans. Image Process.*, vol. 27, no. 2, pp. 703–717, Feb. 2018.

[2] C. Yang, J. Pu, Y. Dong, G.-S. Xie, Y. Si, and Z. Liu, "Scene classification-oriented saliency detection via the modularized prescription," in *Proc. Vis. Comput.*, 2018, pp. 1–16.

[3] T. Alshawi, Z. Long, and G. AlRegib, "Unsupervised uncertainty estimation using spatiotemporal cues in video saliency detection," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2818–2827, Jun. 2018.

[4] K. Oh and K. You, "Visual saliency region detection by combination of soft-and hard-segmentation-wise approaches," *Proc. SPIE*, vol. 27, no. 5, p. 051204, 2018.

[5] T. Chen, M.-M. Cheng, P. Tan, A. Shamir, and S.-M. Hu, "Sketch2photo: Internet image montage," *ACM Trans. Graph.*, vol. 28, no. 5, p. 124, 2009.

[6] E. Vig, M. Dorr, and D. Cox, "Space-variant descriptor sampling for action recognition based on saliency and eye movements," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2012, pp. 84–97.

[7] K. Rapantzikos, Y. Avrithis, and S. Kollias, "Dense saliency-based spatiotemporal feature points for action recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1454–1461.

[8] M.-M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr, and S.-M. Hu, "Global contrast based salient region detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 569–582, Mar. 2015.

[9] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[10] J. Han, H. Chen, N. Liu, C. Yan, and X. Li, "CNNs-based RGB-D saliency detection via cross-view transfer and multiview fusion," *IEEE Trans. Cybern.*, vol. 48, no. 11, pp. 3171–3183, Nov. 2018.

[11] Y. Zhai and M. Shah, "Visual attention detection in video sequences using spatiotemporal cues," in *Proc. 14th ACM Int. Conf. Multimedia*, Oct. 2006, pp. 815–824.

[12] H. Jiang, J. Wang, Z. Yuan, T. Liu, and N. Zheng, "Automatic salient object segmentation based on context and shape prior," in *Proc. BMVC*, 2011, vol. 6, no. 7, p. 9.

[13] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Comput. Vis.*, vol. 59, no. 2, pp. 167–181, Sep. 2004.

[14] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[15] A. Borji, D. N. Sihite, and L. Itti, "Salient object detection: A benchmark," in *Computer Vision*. Berlin, Germany: Springer, 2012, pp. 414–429.

[16] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1597–1604.

[17] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 10, hboxpp. 1915–1926, Oct. 2012.

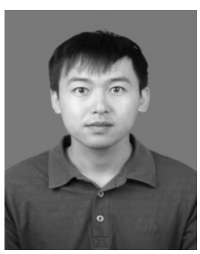
[18] R. Margolin, A. Tal, and L. Zelnik-Manor, "What makes a patch distinct?" in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 1139–1146.

[19] T. Liu *et al.*, "Learning to detect a salient object," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 2, pp. 353–367, Feb. 2011.

[20] S. Alpert, M. Galun, A. Brandt, and R. Basri, "Image segmentation by probabilistic bottom-up aggregation and cue integration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 2, pp. 315–327, Feb. 2012.

[21] V. Movahedi and J. H. Elder, "Design and perceptual validation of performance measures for salient object segmentation. in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2010, pp. 49–56.

[22] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.



JUNFENG WU was born in Dalian, China, in 1983. He received the bachelor's and master's degrees from Liaoning Normal University, in 2006 and 2009, respectively, and the Ph.D. degree from Dalian Maritime University, in 2017. He entered the Postdoctoral Mobile Station with the School of Computer Science and Technology, Tianjin University, as a Researcher, in 2018. He is currently a Lecturer with the School of Information Engineering, Dalian Ocean University, China. His research interests include digital multimedia, data mining, application of digital image processing, and computer vision.

HONG YU was born in Dalian, China, in 1968. She received the bachelor's degree from Jilin University, in 1990, and the master's and Ph.D. degrees from the Dalian University of Technology, in 1999 and 2007, respectively. She is currently a Professor with the School of Information Engineering, Dalian Ocean University, China. Her research interests include digital multimedia, data mining, and big data.

JIANWEI SUN was born in Shandong, China, in 1992. He received the master's degree from Dalian Ocean University, in 2017. He is currently an AI Engineer with Baidu Network Science Company, Beijing. His research interests include image retrieval and data mining.

WENYU QU was born in Yichun, China, in1972. She received the bachelor's and master's degrees from the Dalian University of Technology, in 1994 and 1997, respectively, and the Ph.D. degree from the Japan Advanced Institute of Science and Technology, in 2006. She is currently a Professor with the School of Software, Tianjin University. Her research interests include mobile agent technology, information retrieval, cloud computing, and computer vision.

ZHEN CUI was born in Dandong, China, in 1995. She received the bachelor's degree from Dalian Ocean University, China, in 2017, where she is currently a Graduate Student with the School of Information Engineering. Her research interests include digital multimedia and data mining.