International Journal of Innovative
Computing, Information and Control
Volume **17**, Number **4**, August **2021**

ICIC International ©2021 ISSN 1349-4198
pp. 1435–1446

A MULTI-CUES BASED APPROACH FOR VISUAL SALIENCY DETECTION

Qiaorong Zhang* and Yingfeng Wang

College of Computer and Information Engineering
Henan University of Economics and Law
No. 180, Jinshui East Road, Zhengzhou 450046, P. R. China
*Corresponding author: zqr@huel.edu.cn; wyf96126@126.com

Received February 2021; revised May 2021

Abstract. *Visual saliency detection is very useful in image and video processing. A visual saliency detection approach based on multiple cues is proposed in this paper. Firstly, we preprocess and decompose the original image into super pixels. The feature difference between each super pixel and other super pixels is calculated to get the feature contrast saliency map. Secondly, we extract the background regions from the edge regions of the image and calculate a background based saliency map according to the background regions. Thirdly, we estimate the foreground objects based on the background based saliency map, and calculate the foreground based saliency map. Finally, we integrate the feature contrast saliency map, background based saliency map and foreground based saliency map to get the final saliency map. Evaluations on four benchmark datasets show that the proposed algorithm outperforms other state-of-the-art algorithms on these datasets.*
Keywords: Visual saliency, Saliency map, Background prior, Object estimation

1. **Introduction.** Visual saliency detection is an image processing method developed on the basis of human visual attention mechanism. Through saliency detection, the key areas of the image can be extracted, and the computing resources are allocated preferentially, which can effectively improve the efficiency and accuracy of image processing. Therefore, saliency detection technology is widely used in various visual tasks, such as object detection, object tracking, image segmentation and image retrieval.

In the past decade, visual saliency detection has made remarkable progress, and a large number of methods have been proposed and achieved excellent performance. Saliency detection methods can be classified into two categories: top-down methods and bottom-up methods. Top-down approaches are task driven and achieve high performance by using supervised learning with labels [1]. In particular, deep learning technique has demonstrated the superior ability in saliency detection. Some deep networks for saliency detection have been proposed and have made a qualitative leap in saliency detection performance. However, the deep learning based methods always rely on a large number of labeled data for supervised training. Constructing large-scale labeled training datasets for saliency detection is very laborious and time consuming. At the same time, deep learning methods need high-performance system for training and testing steps, which is difficult to meet in some practical applications [2]. Therefore, this paper focuses on the bottom-up saliency detection method.

Bottom-up saliency is stimulus driven, derived from the salient areas that "pop up" from the visual scene to our eyes. The bottom-up approaches focus on exploring low-level visual features, and detect visual saliency by calculating the differences in visual features of each

DOI: 10.24507/ijicic.17.04.1435

part of the image. Inspired by human visual system, some visual priors are used to describe the characteristics of salient objects, such as contrast prior [3], background prior [4,5] and center prior [6]. Among these methods, the performance of the background prior based method is excellent. However, it is found that the results of background extraction have great influence on saliency detection. False background information will interfere with the detection of salient areas, reduce the detection quality, and make the detection results unstable or fail [7].

This paper proposed the solutions to the above problems. Our first contribution is a novel and reliable background extraction method. Instead of simply regarding the image boundary as the background or the image regions connected with the image boundary as the background, the proposed method assumes that an image region is background only when it is closely connected with the image boundary and has similar features. This method is more robust because it describes the spatial layout of the image region relative to the image boundary and the similarity with the visual features of the image boundary.

Our second contribution is to fuse multiple low-level cues, combining background cues, foreground cues and feature contrast information into a unified formula. The proposed unified mechanism makes use of the advantages of saliency maps based on background, foreground and contrast, and achieves better saliency detection results.

The rest of this paper is organized as follows. Existing visual saliency detection methods are reviewed in Section 2. The proposed multi-priors based approach is described in Section 3. In Section 4, the experiment results and analysis are shown in detail. Finally, the conclusion and future work are summarized in Section 5.

2. **Related Work.** The study of human visual attention mechanism shows that the visual saliency is derived from the uniqueness, scarcity and singularity of visual information. Many existing methods measure feature contrast to compute saliency. For example, Itti et al. [8] proposed a method based on multi-feature and multi-scale center-surround difference. Ma and Zhang [9] presented a fuzzy growth method based on local contrast. Rahtu et al. [10] proposed a Bayesian framework based method. Some researchers calculate visual saliency from a global perspective. Some representative methods include the frequency tuned method proposed by Achanta et al. [11], the global contrast based method proposed by Cheng et al. [12], and saliency filters method proposed by Perazzi et al. [13]. Based on the information theory, some researchers calculate saliency by analyzing the scarcity of features [14,15]. Some researchers have proposed saliency calculation methods based on frequency-domain analysis to improve the computation efficiency, for example, the phase spectrum method proposed by Guo et al. [16], and Hou and Zhang's method of using spectrum residuals [17]. These methods have achieved good results, but there are still some problems: the high saliency values are concentrated on the edges with high contrast, while the saliency values inside the objects are relatively low, and the salient regions are incomplete.

In recent years, saliency detection methods based on graph theory and background prior have attracted great attention. The boundary prior is more popular than center prior which is used previously. Wei et al. [4] used image boundaries as background and computed the shortest path to the background as the saliency. In [5], a saliency detection method based on roust background measures is proposed. Yang et al. [18] presented an approach based on manifold ranking. A saliency detected model based on selective edges prior is proposed in [19]. Xu and Tang [20] exploited hierarchical prior estimation for salient object detection. Wang et al. [21] presented a saliency calculation method via a novel graph model and background. The background based approaches have achieved performances on most images, but there are also some defects. Most methods use the image boundary as the

background. When objects touch the boundary, the result will be deviant. On the other hand, the real background of the image is not only the boundary, but the other regions may also be the background. At the same time, it is not reasonable to consider only the background prior and ignore other priors.

Through the analysis of previous work, this paper studies a visual saliency detection method based on multiple cues, which integrates background prior, object prior and feature contrast prior to improve the accuracy of saliency detection results.

3. **The Proposed Approach.** The visual saliency detection algorithm presented in this paper is described as follows.

Algorithm 1: Visual saliency detection via multi-cues

Input: input image I and related parameters
Output: a saliency map with the same size of the input image
1: Over segment the input image I into n super pixels $S = \{s_1, s_2, \ldots, s_n\}$ using SLIC [22].
2: Compute saliency map SM_{fc} based on feature contrast prior.
3: Construct a weighted undirected graph G.
4: Apply background estimation using three image borders as background seeds.
5: Compute initial saliency map SM_{bg} based on background prior.
6: Apply object estimation using the saliency map SM_{bg}.
7: Compute saliency map SM_{fg} based on foreground object prior.
8: Integrate SM_{fc}, SM_{bg}, and SM_{fg} to obtain a final saliency map.
9: Return the final saliency map.

3.1. **Graph construction.** The input image I is over segmented into n super pixels $S = \{s_1, s_2, \ldots, s_n\}$ using SLIC algorithm [22]. Each super pixel is described using the mean color value of its pixels. Then we construct a weighted undirected graph $G = (V, E)$. V is a set of nodes which are corresponding to the super pixels. E is the edge set. There is an edge between two directly adjacent super pixels.

The adjacent matrix of the graph G is defined as $A = [a_{ij}]_{n*n}$. If the two nodes s_i and s_j are directly adjacent, $a_{ij} = 1$, otherwise $a_{ij} = 0$.

Then, the weighted matrix $W = [w_{ij}]_{n*n}$ is constructed. If the super-pixels s_i and s_j are adjacent, there is an edge e_{ij} between s_i and s_j, whose weight w_{ij} is defined as the feature difference between the super-pixels s_i and s_j. The feature selected in this paper is color feature. If the super pixels s_i and s_j are not adjacent, the weights between s_i and s_j are defined as the length of the shortest path between s_i and s_j. The calculation method of w_{ij} is defined as:

$$w_{ij} = \begin{cases} \left\| S_i^{lab} - S_j^{lab} \right\| & \text{if } a_{ij} = 1 \\ \min\limits_{\rho_1 = S_i, \rho_2 = S_{i+1}, \ldots, \rho_m = S_j} \sum\limits_{k=1}^{m-1} \left\| S_k^{lab} - S_{k+1}^{lab} \right\| & \text{if } a_{ij} = 0 \\ 0 & \text{if } i = j \end{cases} \tag{1}$$

where S_i^{lab} and S_j^{lab} are the mean color values of super pixels s_i and s_j in the CIELab color space.

3.2. **Background estimation.** Background estimation is a key step of the visual saliency calculation method based on background prior. At present, most methods use four boundaries of the image as background. However, the boundary of the image is not always the background. Some boundary super pixels may also be parts of the foreground object.

Meanwhile, other regions which are not boundary of the image may also be the background. The background estimation method used in this paper is based on the following hypothesis.

1) The boundary of the image is usually the background, and the background regions are usually connected to the boundary. This is determined by the characteristics of the human visual system. The human visual system has the highest resolution for the object directly in front of eyes, but can only perceive the outline of the surroundings. Therefore, the objects of interest are generally in the center of the image, and the background is located at the edges of the image. Another reason is that the foreground objects can be placed off the center (one-third rule in professional photography), but they seldom touch the image boundary.

2) There are great differences in color or texture features between the foreground objects and the background regions. The background of an image often has continuous and consistent color or texture features, such as sky, and ocean, while the prominent foreground objects are usually people or other things with different visual characteristics.

3) The interior of the background has smaller differences in color or texture features. This assumption is based on the appearance characteristics of the real world background in images, that is, background areas are usually large and homogeneous. In other words, local or most regions of the background have similar color or texture features.

The steps of background estimation method are described as follows.

Firstly, we choose three borders of the image as background seeds. This is because foreground objects may touch borders of the image. According to the assumptions mentioned above, there is a great difference between the super pixels on the border which contains foreground objects. Therefore, we calculate the standard difference between the color features of super pixels on each border and remove one border with the largest standard difference.

Secondly, background expansion is carried out on the basis of background seeds. For each super pixel which is not in the background set, if it is adjacent to one of the super pixels in the current background and its feature difference is less than the average difference value of the background regions, it will be included in the background set BS.

Figure 1 shows the effect of our background estimation method. The images from left to right are input images, saliency maps using four borders as background and saliency maps based on our background estimation results. It is indicated that our method performs better because it extracts background regions more reasonably.

3.3. **Background based saliency map.** After background estimation and the construction of weighted matrix, the feature difference between each super pixel s_i and background set BS can be calculated according to the weighted matrix W. The larger the difference is, the more significant the super pixel is relative to the background area. We call it background-based saliency. The spatial position between two super pixels is taken into account and the saliency value of each super pixel is normalized to $[0, 1]$. The specific calculation method is defined as

$$\begin{aligned} diff(s_i, BS) &= \sum_{S_j \in BS} weight(s_i, s_j) * \exp\left(-\frac{dist^2(s_i, s_j)}{2\sigma^2}\right) \\ SM_{bg}(s_i) &= \frac{diff(s_i, BS)}{Max(dist(s_i, BS))} \end{aligned} \tag{2}$$

where $dist(s_i, s_j)$ denotes the spatial distance between the center positions of super pixels s_i, and s_j. σ is a constant. In this paper, $\sigma = 0.25$ as in [13] for a balance between local and global effects.

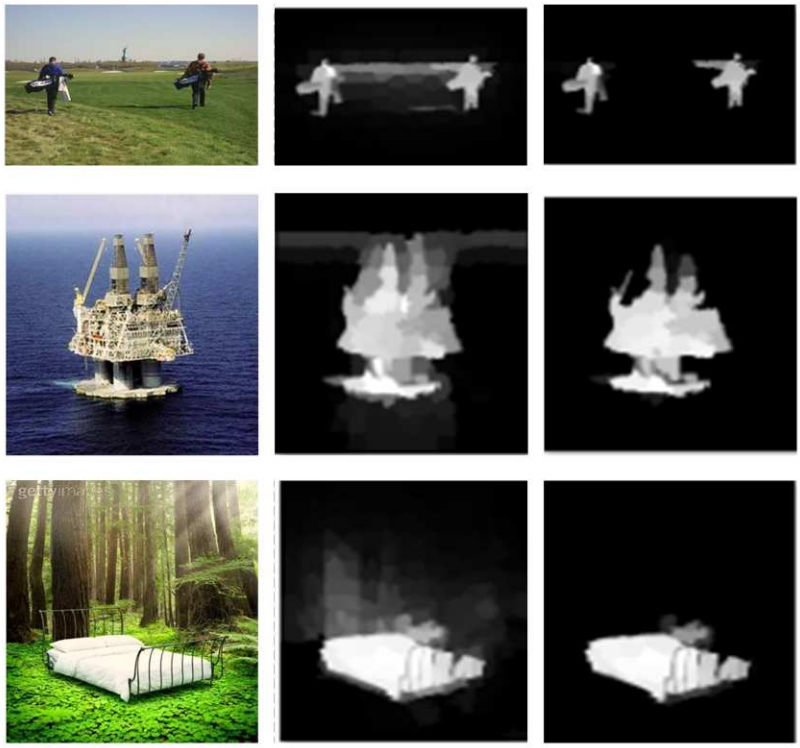

FIGURE 1. Comparison of different backgrounds

3.4. **Foreground based saliency map.** The results of background-based saliency calculation depend on the selection of background area. When the objects touch the edge of the image or the scene of the image is complex, the results of background estimation will be biased, which leads to the inaccurate saliency results. For this reason, we make further optimization on the basis of the background-based saliency calculation results.

According to the background based saliency map, we use adaptive threshold segmentation method to extract the super-pixel region with high saliency value as the foreground area *FS*, and use the similar method of Equation (2) to calculate the feature difference between each super-pixel and the foreground area in the image. Then, Equation (3) is used to calculate the foreground-based saliency of each super pixel.

$$SM_{fg}(s_i) = 1 - \frac{diff(s_i, FS)}{Max(dist(s_i, FS))} \tag{3}$$

3.5. **Feature contrast based saliency.** Those areas (or objects) with novel and distinctive visual features in the image will automatically "pop out" and attract the attention of the observer. The saliency of a region in an image depends on the feature difference between it and other regions. The larger the difference is, the more salient the region is. In addition, the saliency of a region mainly depends on the feature differences with the surrounding regions, and the influence of the regions far away is less. Therefore, we calculate the feature contrast based saliency of each super pixel using Equation (4):

$$SM_{fc}(s_i) = \sum_{s_j \neq s_i} \left\| S_i^{lab} - S_j^{lab} \right\| * \exp\left(-\frac{dist^2(s_i, s_j)}{2\sigma^2} \right) \tag{4}$$

where S_i^{lab} and S_j^{lab} are the mean color values of super pixels s_i and s_j in the CIELab color space. $dist(s_i, s_j)$ denotes the spatial distance between the center positions of super pixels s_i, and s_j.

3.6. **Fusion.** To provide more accurate saliency results, we combine background based saliency map, foreground based saliency map and feature contrast based saliency map using Equation (5):

$$SM(s_i) = w_1 SM_{bg}(S_i) + w_2 SM_{fg}(S_i) + w_3 SM_{fc}(S_i) \tag{5}$$

where w_1, w_2 and w_3 are weights of background based saliency value, foreground based saliency and feature contrast based saliency value. In this paper, we set $w_1 = 0.4$, $w_2 = 0.2$ and $w_3 = 0.4$ according to the best experiment results.

4. **Experiments.** To evaluate the effectiveness of the proposed method, we test our method on four challenging datasets: ECSSD [23], PASCAL-S [24], SOD [25] and DUT-OMRON [18]. These four test datasets provide Ground Truth images labeled manually, which facilitates objective evaluation of the experimental results. ECSSD contains 1000 images with multiple objects, which makes the scene more complex and the detection more difficult. The PASCAL-S dataset contains 850 images, and the selection of these images avoids the subjective bias of intentionally emphasizing the notion of saliency. SOD dataset contains 300 images, which is considered to be a very challenging test dataset. Each image contains one or more objects and structurally complex background. DUT-OMRON consists of 168 high quality images manually selected from more than 140000 images with pixel-wise Ground-Truth labeling [18].

We qualitatively and quantitatively compare our method with eight state-of-the-art methods including GS [4], BD [5], FT [11], RC [12], SF [13], MR [18], WMR [26] and HS [23]. GS, MR, WMR and BD are background-based algorithms. FT, RC and SF are based on feature contrast. HS is a hierarchical saliency detection algorithm. The authors of these algorithms have provided the source codes of the algorithms, which is convenient for us to carry out experimental comparisons.

4.1. **Qualitative evaluation.** Comparison of saliency results between the proposed approach and other eight approaches on four datasets is shown in Figure 2. Intuitively, the results of GS, MR, BD, WMR, HS and our approach are better than those of FT, RC and SF. Compared with GS, MR, BD, WMR and HS, the results of our method are better and closer to the Ground Truth image. Because the proposed approach extracts background more accurately, and combines feature contrast prior and foreground prior, the saliency results are more accurate.

4.2. **Quantitative evaluation.** In the experiments, Precision-Recall curve and F-measure methodologies are used to evaluate the performance of our method and other eight methods. The saliency maps are adjusted to [0 255]. Then the threshold values are selected between 0 and 255 to binaries the saliency maps. Compared with the Ground Truth images, the corresponding Precision and Recall rates are calculated and PR curves are obtained. Precision, Recall and F-measure are calculated using Equation (6):

$$\begin{cases} \text{Precision} = \dfrac{\sum_{x=1}^{M}\sum_{y=1}^{N} GT(x,y) \times B(x,y)}{\sum_{x=1}^{M}\sum_{y=1}^{N} B(x,y)} \\ \text{Recall} = \dfrac{\sum_{x=1}^{M}\sum_{y=1}^{N} GT(x,y) \times B(x,y)}{\sum_{x=1}^{M}\sum_{y=1}^{N} GT(x,y)} \\ \mathrm{F}_\beta = \dfrac{(1+\beta^2)\,\text{Precision} \times \text{Recall}}{\beta^2 \text{Precision} + \text{Recall}} \end{cases} \tag{6}$$

where GT means the Ground Truth image and B represents the binary image after threshold segmentation. The Precision and Recall rates are obtained using an adaptive

FIGURE 2. Comparison of saliency results

threshold which is determined as the mean saliency value of the input image. β is used to determine the importance of Precision and Recall. In this paper, $\beta^2 = 0.3$ makes Precision more influential than Recall.

Comparisons of PR-curve and F-measure are shown in Figure 3 and Figure 4. Generally speaking, the results of background-based visual saliency detection algorithms such as GS, MR, BD, WMR and our algorithm are better than those of FT, RC, SF and other contrast-based detection algorithms. On ECSSD and SOD datasets, the detection results of our

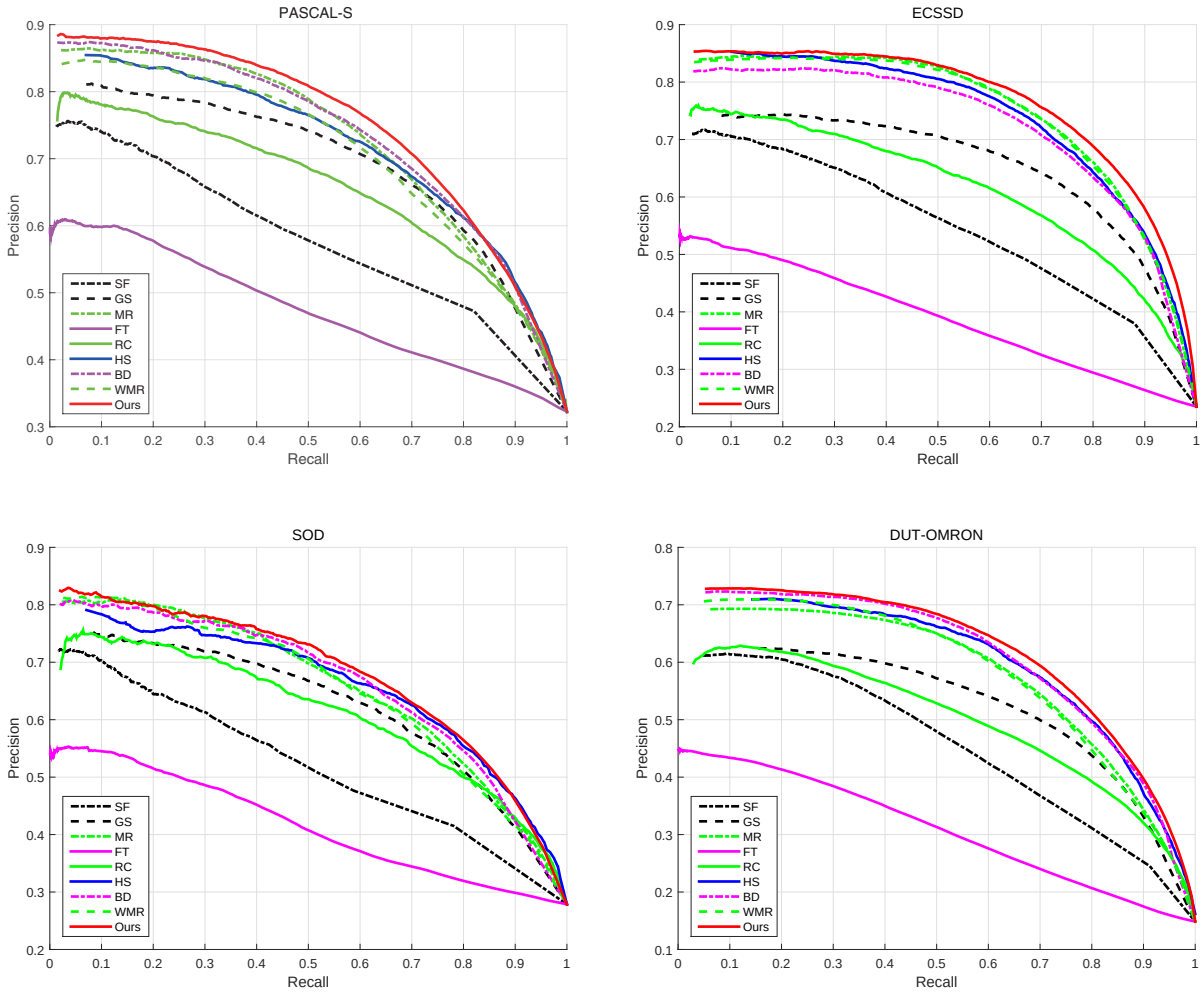

FIGURE 3. PR curve

algorithm are better than those of other eight algorithms. On DUT-OMRON dataset, our algorithm is close to BD algorithm and slightly better than BD algorithm. In our algorithm, a novel and reliable background extraction method is used to describe the spatial layout of the image region relative to the image boundary and the similarity with the visual features of the image boundary. This method has strong robustness. Using this method to extract the background, the saliency map is more accurate. In addition, our algorithm combines background cues, foreground cues and feature contrast cues, makes full use of the advantages of saliency maps based on background, foreground and contrast, and achieves better saliency detection effect.

The Precision, Recall rate and F-measure emphasize the visual saliency of the foreground objects in the image, but neglect the saliency of the background area. Therefore, the visual saliency results cannot be measured comprehensively [20]. MAE (mean absolute error) is another indicator to evaluate the results of visual saliency. Its calculation method is defined as:

$$\mathrm{MAE} = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} |S(x,y) - GT(x,y)| \tag{7}$$

where M and N represent the length and width of the image, $S(x,y)$ is the saliency value of the pixel (x,y), and $GT(x,y)$ is the value of the pixel (x,y) in the Ground Truth image.

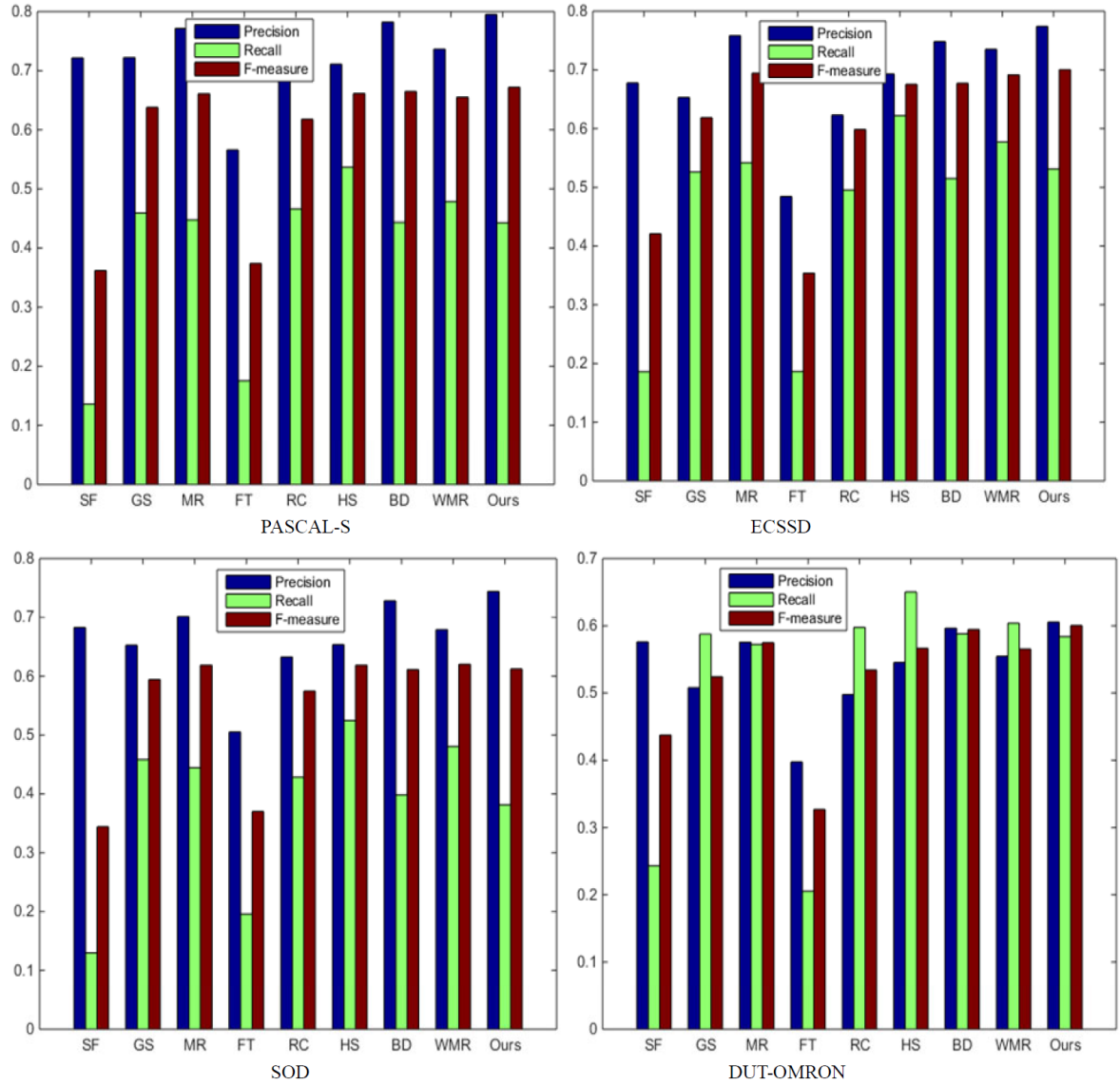

FIGURE 4. F-measure

TABLE 1. Comparison of MAE

	PASCAL-S	ECSSD	SOD	DUT-OMRON
FT	0.2820	0.2696	0.3177	0.2061
RC	0.3117	0.3005	0.3255	0.2903
SF	0.2358	0.2188	0.2689	0.1468
HS	0.2622	0.2275	0.2831	0.2292
GS	0.2209	0.2059	0.2514	0.1732
MR	0.2208	0.1893	0.2593	0.1874
BD	0.1986	0.1714	**0.2296**	0.1438
WMR	0.2341	0.1909	0.2645	0.2008
Ours	**0.1948**	**0.1688**	**0.2296**	**0.1422**

The smaller the MAE value is, the closer the saliency result is to the true value, and the more accurate the result is. The average MAE values of our algorithm and other eight other algorithms on four datasets are shown in Table 1.

From Table 1, we can see that the MAE of our approach is the smallest on DUT-OMRON, ECSSD and PASCAL-S datasets. On SOD dataset, the MAE of our method and BD are the same, less than other approaches. This shows that the saliency detection results of the proposed method are closest to Ground Truth. In addition, in the Precision-Recall curve, the curve of the proposed algorithm is close to that of MR and BD, but the average MAE value of the proposed algorithm is smaller than that of the two algorithms. The reason is that the background area extracted by our algorithm is more accurate and reasonable, so that the saliency of the background area can be suppressed when calculating saliency. Therefore, the MAE obtained is the smallest and the saliency detection result is closer to Ground Truth image.

4.3. **Parameter selection.** The proposed algorithm needs to set three parameters when running, which are the weights w_1, w_2 and w_3 in Formula (5). The curves of Recall, Precision and F-measure corresponding to different values of these three parameters are shown in Figure 5. It can be seen from the figure that when the value of w_1 is in the range of $[0.2, 0.4]$, the value of w_2 is 0.2, and the value of w_3 is in the range of $[0.2, 0.4]$, the value of F-measure remains near the highest value. Therefore, in the experiment we set $w_1 = 0.4$, $w_2 = 0.2$, $w_3 = 0.4$.

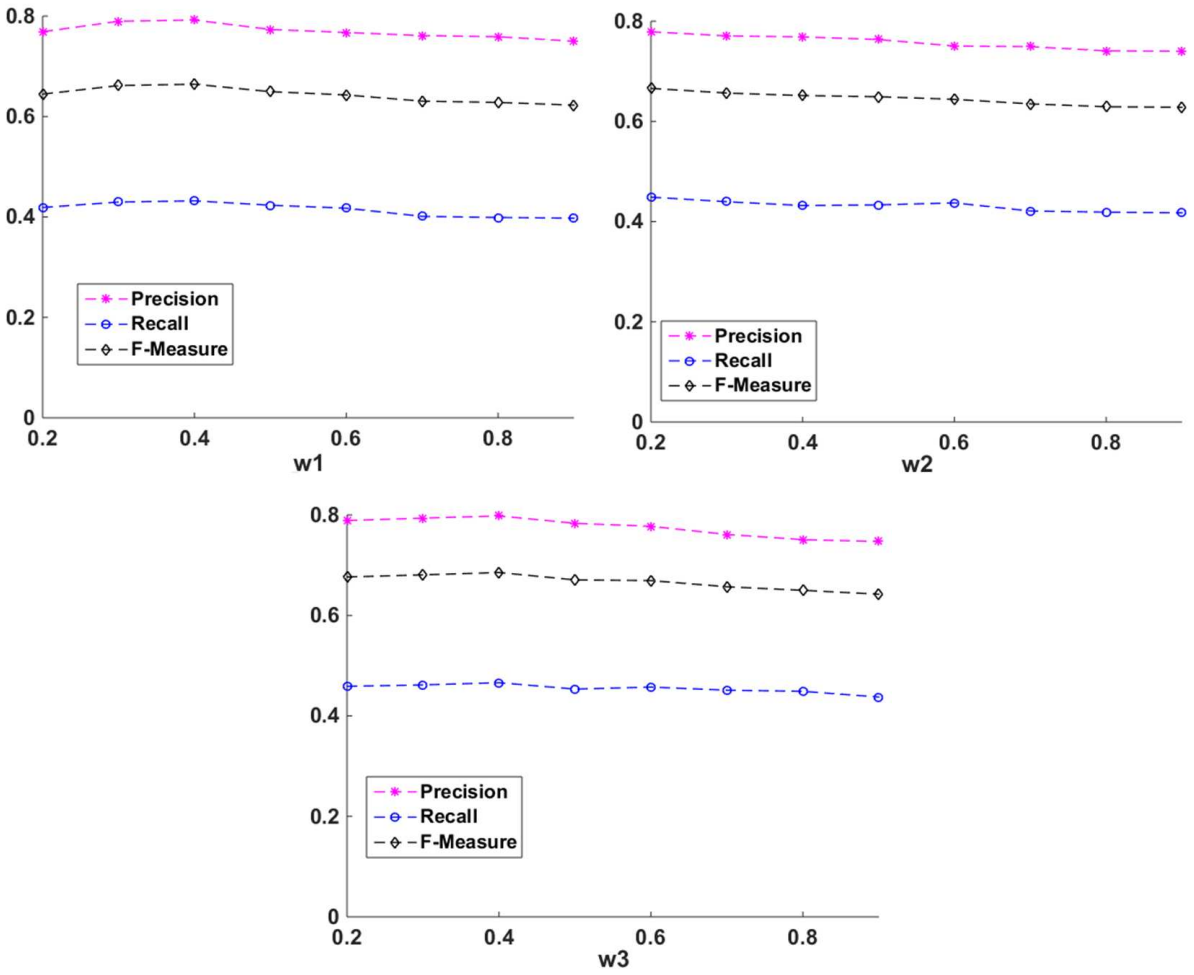

FIGURE 5. Precision, Recall and F-measure change curve

5. **Conclusions.** In this paper, a visual saliency detection approach based on multiple priors is proposed. It can extract background more accurately according to boundary connectivity, feature similarity between background super pixels, and feature difference between background and objects. By fusing background prior, foreground prior and feature contrast prior, the result of visual saliency calculation is more accurate. Experiments on DUT-OMRON, SOD, PASCAL-S and ECSSD are carried out and compared to Ground Truth. Quantitative evaluations on these four datasets show that the proposed algorithm obtains higher Precision, Recall and F-measure values and lower MAE values compared to other eight state-of-the-art methods.

In the experiments, we find that the proposed algorithm and other algorithms fail to detect salient regions when the foreground objects and background area are very similar in color and texture. Therefore, our future work is to consider combining other prior knowledge into the saliency calculation, extracting higher level features and then applying them to saliency detection, improving and optimizing saliency results.

Acknowledgment. This work is partially supported by the Key Scientific Research Project of Henan Higher Education Institutions (No. 18B520005). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] I. A. Siradjuddin, A. Sakinah and M. K. Sophan, Combination of feature engineering and feature learning approaches for classification on visual complexity images, *International Journal of Innovative Computing, Information and Control*, vol.17, no.3, pp.991-1005, 2021.

[2] E. P. Putra, S. Michael, T. O. Wingardi, R. L. Tatulus and W. Budiharto, Smart traffic light model using deep learning and computer vision, *ICIC Express Letters*, vol.15, no.3, pp.297-305, 2021.

[3] M. Cheng, G. Zhang, N. J. Mitra, X. Huang and S. Hu, Global contrast based salient region detection, *Proc. of the 2011 IEEE International Conference on Computer Vision and Pattern Recognition*, Colorado Springs, USA, pp.409-416, 2011.

[4] Y. Wei, F. Wen and W. Zhu, Geodesic saliency using back ground priors, *Proc. of the European Conference on Computer Vision 2012: Part III*, Florence, Italy, pp.29-42, 2012.

[5] W. Zhu, S. Liang and Y. Wei, Saliency optimization from robust background detection, *Proc. of the 2014 IEEE International Conference on Computer Vision and Pattern Recognition*, Columbus, USA, pp.2814-2821, 2014.

[6] C. Yang, L. Zhang and H. Lu, Graph-regularized saliency detection with convex-hull-based center prior, *IEEE Signal Processing Letters*, vol.20, no.7, pp.637-640, 2013.

[7] J. Wu, H. Yu, J. Sun, W. Qu and Z. Cui, Efficient visual saliency detection via multi-cues, *IEEE Access*, no.7, pp.14728-14735, 2019.

[8] L. Itti, C. Kouch and E. Niebur, A model of saliency-based visual attention for rapid scene analysis, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.20, no.11, pp.1254-1259, 1998.

[9] Y. Ma and H. Zhang, Contrast-based image attention analysis by using fuzzy growing, *Proc. of the 11th ACM International Conference on Multimedia*, New York, NY, USA, pp.374-381, 2003.

[10] E. Rahtu, J. Kannala, M. Salo and J. HeikkilÄa, Segmenting salient objects from images and videos, *Proc. of the 11th European Conference on Computer Vision*, Heraklion, Greece, pp.366-379, 2010.

[11] R. Achanta, S. Hemami, F. Estrada and S. Süsstrunk, Frequency-tuned salient region detection, *Proc. of IEEE International Conference on Computer Vision and Pattern Recognition*, Miami, USA, pp.1597-1604, 2009.

[12] M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr and S. Hu, Global contrast based salient region detection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.37, no.3, pp.569-582, 2015.

[13] F. Perazzi, P. KrÄahenbÄuhl, Y. Pritch and A. Hornung, Saliency filters: Contrast based filtering for salient region detection, *Proc. of the 2012 IEEE International Conference on Computer Vision and Pattern Recognition*, Providence, USA, pp.733-740, 2012.

[14] N. Bruce and J. Tsotsos, Saliency based on information maximization, *Proc. of Advances in Neural Information Processing Systems*, Vancouver, British Columbia, pp.5-8, 2005.

[15] N. Bruce and J. Tsotsos, Saliency, attention, and visual search: An information theoretic approach, *Journal of Vision*, vol.9, no.3, pp.1-24, 2009.
[16] C. Guo, Q. Ma and L. Zhang, Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform, *Proc. of the 2008 IEEE Conference on Computer Vision and Pattern Recognition*, Minneapoils, USA, pp.2908-2915, 2008.
[17] X. Hou and L. Zhang, Saliency detection: A spectral residual approach, *Proc. of the 2007 IEEE International Conference on Computer Vision and Pattern Recognition*, Minneapoils, USA, pp.1-8, 2007.
[18] C. Yang, L. Zhang, H. Lu et al., Saliency detection via graph-based manifold ranking, *Proc. of the 2013 IEEE International Conference on Computer Vision and Pattern Recognition*, Portland, OR, USA, pp.3166-3173, 2013.
[19] Y. Jiang, L. Tan and S. Wang, Saliency detected model based on selective edges prior, *Journal of Electronics & Information Technology*, vol.37, no.1, pp.130-136, 2015.
[20] W. Xu and Z. Tang, Exploiting hierarchical prior estimation for salient object detection, *Acta Automatic Sinica*, vol.41, no.4, pp.799-812, 2015.
[21] Q. Wang, W. Zheng and R. Piramuthu, GraB: Visual saliency via novel graph model and background priors, *Proc. of the 2016 IEEE International Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, pp.535-543, 2016.
[22] R. Achanta, A. Shaji and K. Smith, SLIC superpixels compared to state-of-the-art superpixel methods, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.34, no.11, pp.2274-2282, 2012.
[23] Q. Yan, L. Xu, J. Shi et al., Hierarchical saliency detection, *Proc. of the 2013 IEEE International Conference on Computer Vision and Pattern Recognition*, Portland, OR, USA, pp.1155-1162, 2013.
[24] A. Borji, What is a salient object? A dataset and a baseline model for salient object detection, *IEEE Trans. Image Processing*, vol.24, no.2, pp.742-756, 2015.
[25] V. Movahedi and J. H. Elder, Design and perceptual validation of performance measures for salient object segmentation, *Proc. of the 7th IEEE Computer Society Workshop on Perceptual Organization in Computer Vision*, San Francisco, CA, USA, pp.49-56, 2010.
[26] X. Zhu, C. Tang, P. Wang et al., Saliency detection via affinity graph learning and weighted manifold ranking, *Neurocomputing*, vol.312, no.10, pp.239-250, 2018.