



Saliency detection using multiple low-level priors and a propagation mechanism

Muwei Jian^{1,2} · Jing Wang² · Junyu Dong² · Chaoran Cui¹ · Xiushan Nie¹ · Yilong Yin³

Received: 4 July 2018 / Revised: 14 March 2019 / Accepted: 31 May 2019 /

Published online: 11 June 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

The majority of existing methods for saliency detection based on low-level features failed to uniformly highlight the salient-object regions. In order to improve the accuracy and consistency of generated saliency maps, we propose a novel and efficient framework by combining low-level saliency priors and local similarity cues for image saliency detection. Firstly, we construct a multiple low-level prior map using location prior, color prior and background prior. Then, the prior maps employ a propagation mechanism based on Cellular Automata to enforce relevance of similar regions as a local similarity cue. Finally, a principle refinement framework by integrating multi-level prior maps and local similarity cue map are used to obtain an ultimate high-quality saliency map. Extensive experiments on publicly available datasets show that our designed approach is capable of producing accurate saliency maps compared with those generated results by the state-of-the-art saliency-detection methods.

Keywords Multimedia data analysis · Background prior · Saliency detection · Cellular Automata · Multimedia data modeling

1 Introduction

Visual saliency detection (VSD) is a computational model that simulates the human visual system (HVS), which aims to identify important regions or objects in an image or video. In the past few decades, it has been applied to computer vision and neurosciences, including image segmentation [3, 19], image matching [38], video tracking [22, 27, 28], visual relation detection [37, 46], event-based multimedia analysis [23, 25], etc.

✉ Muwei Jian
jianmuweihk@163.com

¹ School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan, China

² Department of Computer Science and Technology, Ocean University of China, Qingdao, China

³ School of Software Engineering, Shandong University, Jinan 250101, China

According to whether prior knowledge is used or not, the current saliency-detection models can be divided into two categories: bottom-up and top-down ones. Bottom-up models [1, 4, 7, 8, 10, 11, 20, 21, 24, 26, 39, 45, 47] are stimulus-driven, using low-level features, such as color, texture, contrast and spatial location cues. On the other hand, top-down based models usually exploit high-level human perceptual knowledge, semantics and object priors to guide saliency computation [6, 9, 42], which are often task-driven and special application-oriented with huge training data and prior knowledge.

It was commonly accepted that the efficacious integration of multiple low-level cues can produce reliable and robust saliency-detection results [14]. Thus, in this paper, we obtain preliminary saliency maps by integrating multiple layers of prior knowledge. Then we exploited the local and global contrast based on the colour and texture feature of the image region [4, 26] to calculate a tentative saliency map. Since human beings are subjective when looking at an image, the simple contrast calculation may not be enough to provide adequate visual clues. In fact, human beings are accustomed to focus on more colourful areas in the center of the image, whereas the dark areas in the boundary of the image are easily ignored. Therefore, we adopt three low-level prior-knowledge cues to generate an initial saliency map, including colour prior and spatial location prior and background prior. Then, we refine the preliminary saliency map based on Cellular Automata (CA).

In contrast to previous works, in this paper, we propose an effective and compound saliency-detection model, which is based on multi-level prior and a propagation mechanism relying on Cellular Automata. In our method, we firstly divide the input image into compact uniform elements. Then, we incorporate the Convex-Hull-Based center method to obtain the spatial location of the salient objects. According to psychology research, the warm colors such as red and yellow are more remarkable and distinct for human visual perception. In order to use this fundamental information, we also exploit color prior during the procedure of saliency detection. Meanwhile, in order to capture the boundary details of salient object, we generate the background-prior map by calculating the probability that the image region is connected to the image boundaries. Finally, these three typical maps are integrated to obtain a multi-level saliency map, and then we apply a propagation mechanism based on CA to reduce the redundant information in similar regions to obtain a refined saliency map for salient-object detection. In order to evaluate the proposed saliency-detection model, we carried out extensive experiments on several widely used datasets and experimental results show that the proposed method outperforms other state-of-the-art saliency-detection approaches.

The rest of this paper is organized as follows. In Section 2, we review the related work on saliency detection models. In Section 3, we describe the proposed framework in detail. Section 4 shows the experimental results. Finally, conclusion and discussion are presented in Section 5.

2 Related work

Most of the existing visual saliency-detection (VSD) methods are based on the bottom-up computational scheme, which uses low-level visual features in images or videos. Early work of saliency detection was proposed by Itti et al. [13]. In this model, based on visual characteristics (include colour, intensity and orientation maps), center-surrounding difference of low-level image features were used to obtain the final saliency map of an input

image. In [26], Ma et al. proposed an approach for visual attention area detection in images based on local contrast analysis, which was computed based on center-surrounding and colour difference. Harel and Perona [8] proposed a graph-based visual saliency-detection (GBVS) model aiming to fuse different feature maps so as to strengthen those salient regions in an image. However, this method produces a lower resolution saliency map. Liu et al. [24] used Conditional Random Fields (CRF) to identify salient region in an image, which employed three types of low-level features including the multi-resolution contrast, the center-surround histogram, and the colour spatial distribution. Hou et al. [11] proposed a spectral-residual (SR) saliency-detection approach by analysing the log-spectrum of an image, with the intent of focusing on the contrast while ignoring the redundant information. In [1], a frequency-tuned (FT) based model for saliency detection was proposed by calculating the colour difference from the mean image colour at the pixel level. Goferman et al. [7] proposed a context-aware (CA) saliency-detection method based on local contrast and human visual attention rules to detect salient objects. In [45], Zhai and Shah proposed a fusion spatial attention model and temporal attention model for video saliency detection. In [20], a high-dimensional colour transform method was proposed for salient object detection. Cheng et al. [4] proposed a histogram-based contrast (HC) and a region-based contrast (RC) method for saliency detection. Li et al. [21] proposed a salient object detection model by combing meanshift filtering and fusing of colour information in an image. In [47], Zhang and Sclaroff designed a boolean-map based method by exploiting surroundedness for saliency detection. Wei et al. [39] exploited background priors and geodesic distance for saliency detection. Zhu et al. [48] integrated boundary connectivity and multiple low-level cues to estimate saliency of images. Shen et al. [34] proposed a saliency-detection model via unifying high-level priors into a low-rank framework. In [31], a novel structured matrix decomposition model (SMD) with two structural regularizations, namely tree-structural sparsity-inducing regularizations and Laplacian regularizations, was developed for salient object detection. Qin et al. [32] proposed a model by using global distance matrices and applied Cellular Automata to optimize the prior map for salient object detection. In [44], Yang, et al. proposed a graph-based manifold ranking method, which used boundary regions consider both foreground and background cues. Song et al. [36] proposed a co-saliency method based on depth cues to detect the salient object objects. In [12], by taking the proposals as the bags of instances of multiple instances learning (MIL), where the instances were the superpixels contained in the proposals, Huang et al. formulated saliency detection problem as a MIL task. Chang et al. [3] proposed a novel computational framework for saliency detection with the aid of minimizing energy function and knowledge prior. In [17], a salient object segmentation method was proposed by integrating both bottom-up salient stimuli and object-level shape prior. Based on a set of representative texture atoms, Scharfenberger et al. [33] proposed a sparse texture model for the saliency object detection through extracting rotation-invariant neighborhood-based textural representations. In [41], Yan et al. proposed a multi-layer approach to analyse saliency cues for saliency detection. Cheng et al. [5] proposed a method using a soft image abstraction representation to decompose an image into large scale perceptually homogeneous elements for efficient salient region detection. Recently, a saliency detection method, using the quaternionic distance based weber local descriptor and low-level priors, is proposed for natural images [16]. In [15], Jian et al. proposed a saliency-detection model by integrating quaternionic distance based weber descriptor (QDWD) with pattern distinctness and local contrast for underwater vision

saliency detection. In [18], a novel computational model for saliency detection is proposed by combining the local and global kernel density estimations during the saliency computation process.

Compared to bottom-up saliency-detection model, top-down methods are based on the high-level priors of images and therefore is always specific task-driven with training and learning scheme. In [42], Yang et al. proposed a top-down saliency detection model by jointly learning a Conditional Random Field (CRF) and a visual dictionary for salient object detection. Cholakkal et al. [16] designed a top-down saliency detection method based on locality-constrained contextual sparse coding. In their approach, three locality constraints were combined to label an object category. In [15], a two-stage deep method was exploited to learn the intra-class association between the exemplars and query objects for top-down saliency detection.

In this paper, we concentrate on the popular category of the bottom-up computational saliency detection model. Different from conventional approaches, the proposed scheme can ameliorate the accuracy and consistency of the estimated saliency results by fusing multi-level saliency priors and local similarity cues. The main contributions of the proposed approach are twofold:

- An efficient and compound framework by integrating low-level saliency priors and local similarity cue is proposed for visual saliency detection.
- With the merit of propagation mechanism based on CA, a principle refinement framework is devised to combine multi-level prior maps and local similarity cues together for uniformly highlighting salient objects. As a result, the scheme can reduce the difference among similar regions and suppress the background by ameliorating the multi-level prior maps.

3 The proposed method

In this section, we describe the designed saliency-detection framework, which incorporates multiple low-level prior maps together through a propagation mechanism based on Cellular Automata and saliency refinement procedure. Multi-level priors will first be described, including center prior, colour prior and background prior. All these different types of information are fused to obtain an integrated saliency map.

3.1 Multi-level prior Integration

Given an input image, we use the simple linear iterative clustering (SLIC) algorithm [2] to segment the image into N superpixels $P = \{p_1, p_2, \dots, p_N\}$.

3.1.1 Center-prior

Plenty of existing center prior methods directly use the image center [7, 24, 26]. Actually, when the salient objects are in the image center, these methods will produce good results. However, there are many salient objects that deviate from the center of image in real situation; in this case these methods are no longer applicable. In order to tackle this issue, in this paper, we adopt the convex hull of interest points to estimate the real spatial center of the salient

object [43]. Firstly, linear iterative clustering (SLIC) algorithm [2] is applied to segment the image into N superpixels, and then compute a convex hull which encloses interesting points to orientate the location of salient object. Secondly, we regard the centroid of the convex hull as the center to achieve the convex-hull-based center prior estimation. The center-prior saliency value of a pixel can be defined as follows:

$$CenP_i = \exp\left(-\frac{\|x_i - x_0\|^2}{2\theta_x^2} - \frac{\|y_i - y_0\|^2}{2\theta_y^2}\right) \quad (1)$$

where x_i and y_i denote the mean horizontal and vertical positions of the superpixel i , (x_0, y_0) denote the center of the superpixel i , θ_x and θ_y indicate the horizontal and vertical variances. We use a centered anisotropic Gaussian distribution to model the center prior map and set $\theta_x = \theta_y$ with pixel coordinate normalize to $[0, 1]$. The convex hull provides a rough location of the salient object, and the convex-hull-based center prior map is more reasonable and robust, which achieves a good performance. Some examples of center prior map are illustrated in Fig. 1. It shows that the convex-hull-based center prior is able to localize the spatial position of the salient object. Even if the object is at the edge of the image, it can still achieve satisfactory results.

3.1.2 Colour-prior

According to the human vision system, human eyes are more sensitive to warm colour such as red and yellow colour. Warm colours are thus more attention-getting for human perception. To exploit this information, we used colour-prior proposed in [34] and obtain a 2-D histogram distribution $H(S)$ in $nR-nG$ colour space, which can be expressed as follows:

$$nR = \frac{R}{R + G + B} \quad (2)$$

$$nG = \frac{G}{R + G + B} \quad (3)$$

Similarly, a histogram indicating the colour distribution of the background $H(\text{背})$ is generated as well. For each quantized colour, we obtain the values from two histograms, denoted by h_s and h_B , respectively. Then, the color prior for saliency estimation can be expressed as follows:

$$ColP_i = \exp\left(\frac{(h_S(c_x) - h_B(c_x))}{\lambda_3^2}\right) \quad (4)$$

where c_x indicates the color at location x . Figure 2c shows some samples of color prior maps.

3.1.3 Background-prior

Recently, some research works [39, 44] introduce boundary prior and treat image boundary regions as background. In [39], the boundary superpixels are used as background queries. While in [44], the authors calculate the shortest-path distance to the image boundary for saliency detection via graph-based manifold ranking. Inspired by these previous works, we

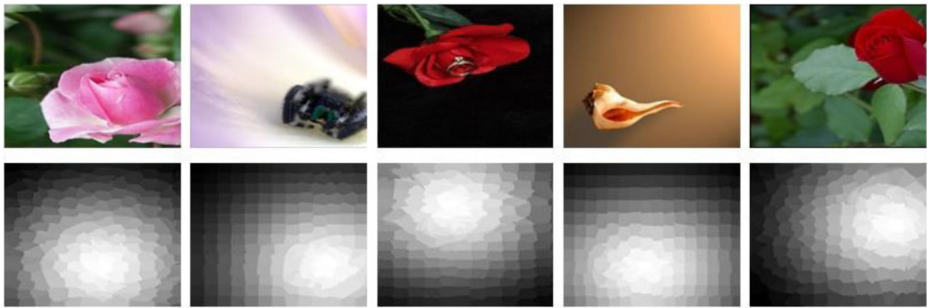


Fig. 1 Examples of center prior maps. The first line represents the input image; the second line is the center-prior cues based on the centroid of the convex hull

calculate the probabilities of image regions connected to image boundaries [48]. The background-prior can be expressed as below:

$$BgP_i = \exp\left(-\frac{bdcon^2(p_i)}{2\sigma^2}\right), \tag{5}$$

where *bdcon* denote the boundary connectivity, parameter $\sigma=0.3$. The background-prior for saliency detection is shown in Fig. 1d.

3.1.4 Multi-level prior fusion

To enhance saliency map, we integrate these three low-level visual cues simultaneously, namely binding color prior, center prior and background prior to obtain multiple low-level prior $S_{prior(i)}$ as follows:

$$S_{prior(i)} = CenP_i.*ColP_i.*BgP_i. \tag{6}$$

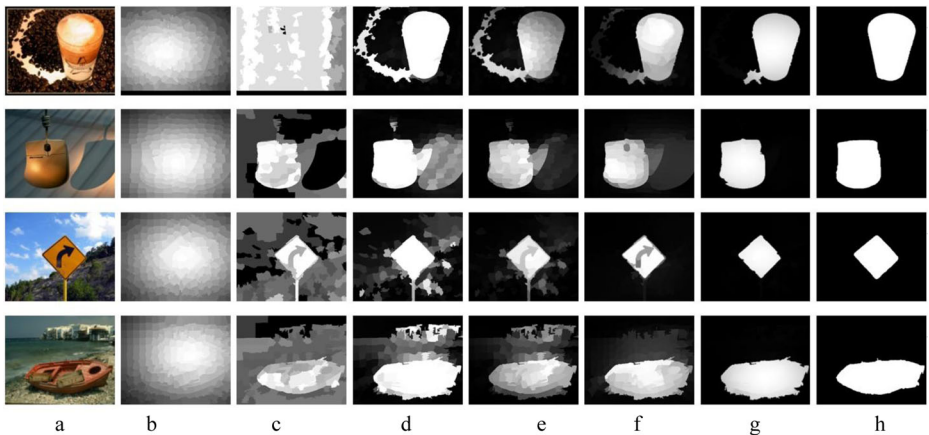


Fig. 2 Samples of multiple low-level prior maps. **a** Initial input images; **b** center prior maps; **c** color prior maps; **d** background prior maps; **e** Fusion results of prior maps; **f** Saliency maps of propagation mechanism; **g** Saliency refinement; and **h** Ground truths

where $\cdot*$ is the pixel-by-pixel multiplication operator. Figure 2e displays some samples of multiple low-level priors.

3.2 Propagation mechanism

In order to strengthen the relevance of similar regions, we employ a propagation mechanism based on Cellular Automata [32, 40]. We detect the neighbors of the superpixel as well as sharing common boundaries with its adjacent superpixels as a cell. If two adjacent image superpixels are similar, we utilize the affinity matrix to represent the similarity of superpixels pair (p_i, p_j) . The affinity matrix $W = (w_{i,j} \in R^{N \times N})$, with $w_{i,j}$ is defined as:

$$w_{i,j} = \begin{cases} \exp\left(-\frac{\|p_i - p_j\|^2}{2\lambda^2}\right), & \text{if } (p_i, p_j) \in V \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

where p_i, p_j denotes superpixel i and j , V denotes the set of adjacent superpixel pairs which are either neighbors on the image. In order to normalize affinity matrix, a degree matrix $D = \text{diag} \{d_1, d_2, \dots, d_N\}$ is generated, where $d_i = \sum_j w_{i,j}$, to form a row-normalized impact affinity matrix [32], which can be written as blow:

$$W^* = D^{-1} \cdot W \tag{8}$$

Considering that each superpixel of the next state is determined by itself and neighbours, we need to balance the importance of the two factors relationship. If a superpixel is quite different from all neighbours in the color space, the next state will primarily rely on itself. On the contrary, if a superpixel is similar to neighbors, it is more likely to be assimilated by the local environment. So we build a coherence matrix $C = \text{diag} \{c_1, c_2, \dots, c_N\}$ to better promote the evolution of all neighbours, each superpixel coherence towards its current state $c_i = 1 / \max(w_{i,j})$, then we can construct the coherence matrix $C^* = \text{diag} \{c_1^*, c_2^*, \dots, c_N^*\}$ by the formulation as in [32]:

$$c_i^* = a \cdot \frac{c_i - \min(c_j)}{\max(c_j) - \min(c_j)} + b \tag{9}$$

where $j = 1, 2, \dots, N$, and set $a = 0.6, b = 0.2$. With coherence matrix C^* , each superpixel can evolve into a more accurate and stable state. Under the influence of neighbouring superpixels, the salient objects can be more easily detected. According to the affinity matrix and coherence matrix, it can be defined the propagation mechanism [32] as follows:

$$S_c = C^* \cdot S^p + (I - C^*) \cdot W^* \cdot S^p \tag{10}$$

where $S^p = \{S_{prior(1)}, S_{prior(2)}, \dots, S_{prior(i)}, \dots, S_{prior(N)}\}$, I is an identity matrix, C^* and W^* represent coherence matrix and affinity matrix, respectively. Figure 2f shows some saliency maps of propagation mechanism.

We also evaluated the performance of the propagation mechanism based on and without CA. As shown in the Fig. 3, the first line represents the input images, while the second line is the saliency map without CA, and the third line shows the saliency maps with CA. By using the CA processing, the effect of image background noises can be reduced, and meanwhile salient object in the image can be highlighted more uniformly. Therefore, salient maps produced by CA contain a clear outlines, so the cellular automata (CA) plays an important role in our devised scheme.

Furthermore, the proposed updating rule applies to most images. Through exploiting the intrinsic relationship in the neighborhood superpixels, our CA-based model can enhance saliency consistency among similar regions. As shown in Fig. 3, CA can suppress the background and uniformly highlight the salient object in an image. In addition, even if the image is at the boundary of the image, such as the first image and the fourth image, satisfactory results can also be obtained.

3.3 Saliency refinement

To obtain a high-quality saliency map, we finally combine multiple low-level saliency prior and propagation mechanism together, and the saliency refinement procedure can be expressed as follows:

$$S_{(i)} = \frac{(M + S_I^{P*} + S_{c(i)*})}{S_{(i)}^{P*}} \quad (11)$$

where M is a Laplacian matrix, S_I^{P*} denotes the value of multiple low-level prior diagonal matrix, $S_{c(i)*}$ denotes the diagonal matrix value of S_c . Figure 2g shows some results of saliency refinement, compared with the ground truths in Fig. 2g, saliency refinement is capable of further enhancing the saliency maps.

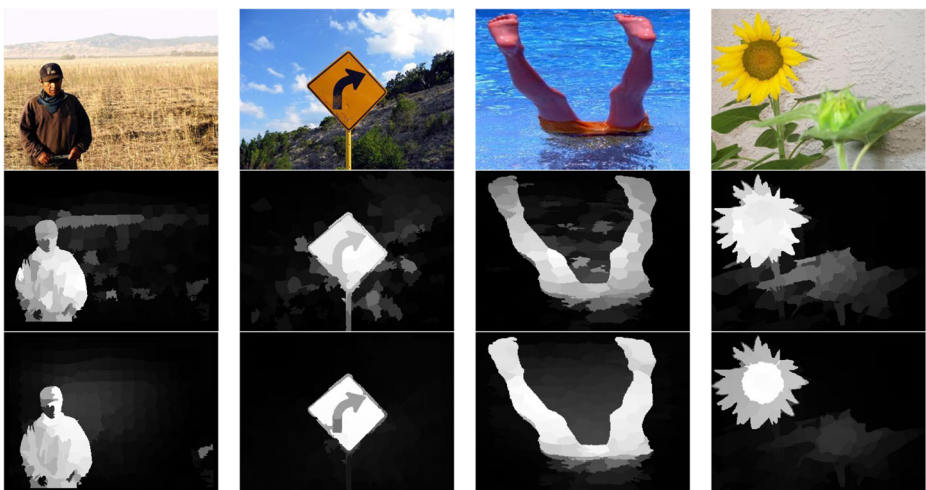


Fig. 3 Comparison experiments by using CA and without CA. The first line are the input images, the second line is the saliency maps without CA, while the third line shows the saliency maps with CA

4 Experimental results

In order to evaluate the performance of our algorithm, we conduct a series of experiments using three widely used datasets, including MSRA [30], ECCSD [35], SOD [29]. The proposed saliency-detection method is compared with other state-of-the-art methods, i.e. Frequency-tuned: FT [1], Robust background detection: RBD [48], Spectral residual: SR [11], Global contrast: RC [4], Manifold ranking: MR [44], Local contrast: LC [4], Histogram contrast: HC [4]. For comprehensive evaluation, we use precision-recall (P-R), F -measure to assess the quality of the saliency maps. Precision is related to the saliency detection correctly of the computation model, Recall is the ration of correctly detection to ground-truth, F -measure is a performance measure harmonic mean of precision (P) and recall (R) [24] quantitatively, which is defined as follows:

$$F_{\beta} = \frac{(1 + \beta^2) \times P \times R}{\beta^2 \times P + R} \tag{12}$$

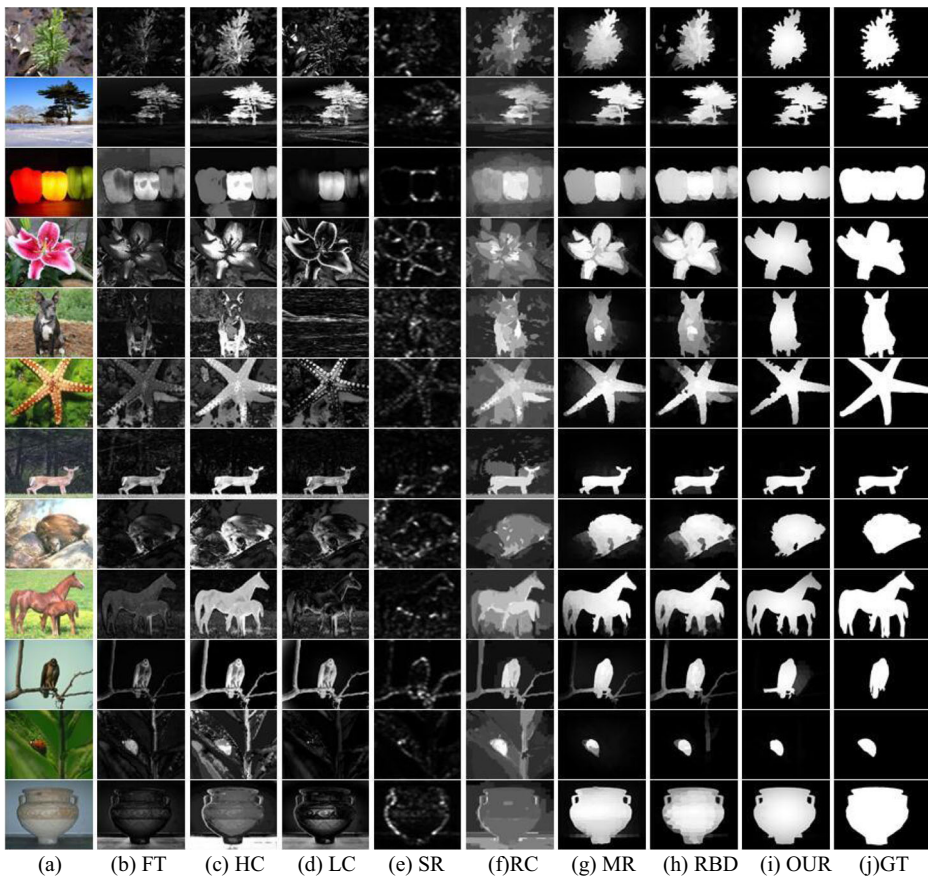


Fig. 4 Comparison of different saliency-detection methods. **a** Input images. **b-h** saliency maps generated by FT [1], HC [4], LC [4], SR [11], RC [4], MR [44], RBD [48], accordingly. **i** Saliency maps of our method. **j** Ground truth. The first to four rows are from MASK dataset, and five to eight rows are from ECCSD dataset, while nine to twelve rows are from SOD dataset, respectively

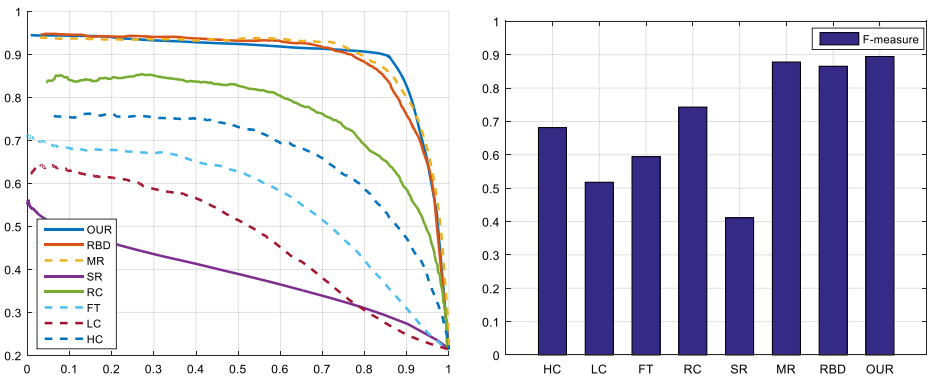


Fig. 5 Comparison of different saliency-detection methods in terms of average precision, recall, and *F*-measure. Left: P-R curve on MSAR-5000 dataset; Right: *F*-measure on MSAR-5000 dataset

where β^2 is set to 0.3 as suggested in [1]. The P-R curve is an important indicator of the model performance. The larger the area under the curve, the more obvious the superiority and robust of the algorithm.

4.1 Evaluation on MSAR 5000 dataset

Firstly, we evaluated our method on the MSAR 5000 dataset for each image with a single salient object and compare our results with other seven methods, including FT, LC, RC, HC, RBD, MR and SR. The visual comparison is given in Fig. 4, it can be seen that the proposed method is able to produce reliable and high quality saliency-detection results. Figure 5 shows the precision, recall and the *F*-measure values of all the different methods. From the comparison, our method can achieve the top precision. In addition, the proposed model outperforms the other existing methods in terms of detection accuracy. Experiment results performed on MSAR 5000 database show that our model is efficient and accurate.

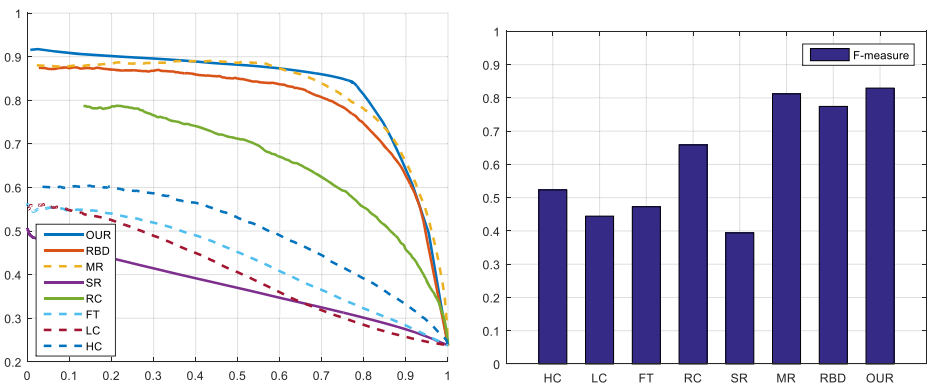


Fig. 6 Comparison of different saliency-detection methods in terms of average precision, recall, and *F*-measure. Left: P-R curve on ECSSD dataset; Right: *F*-measure on ECSSD dataset

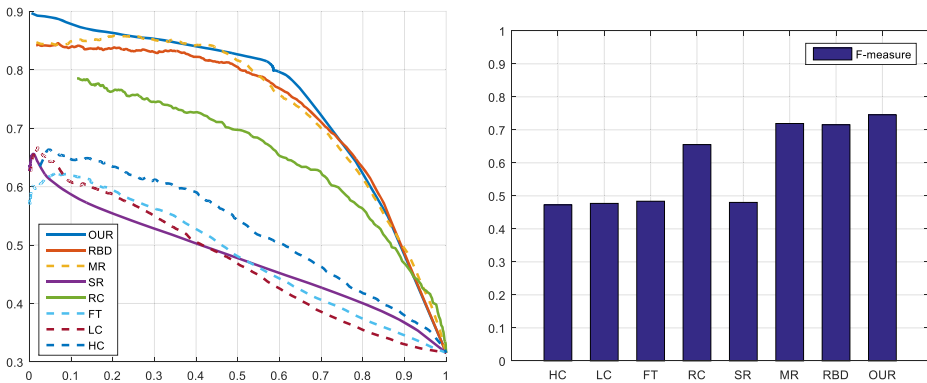


Fig. 7 Comparison of different saliency-detection methods in terms of average precision, recall, and F -measure. Left: P-R curve on SOD dataset; Right: F -measure on SOD dataset

4.2 Evaluation on ECSSD dataset

We also tested our method on the publicly available saliency-detection dataset ECSSD, which contains 1000 images with complex foreground and background. We also quantitatively compared our method with other methods of FT, LC, RC, HC, RBD, MR and SR. Figure 4 (five to eight rows) shows some results of saliency maps generated by different methods. It can be observed that the proposed method can handle complex foreground and background, generating the detection results of almost entire salient objects and preserving the outline of salient object more clearly.

Figure 6 gives the precision, recall and the F -measure values on the dataset ECSSD of individual algorithms. According to P-R curve, the proposed model produces the optimal saliency-detection result. In terms of F -measure, the proposed method also yields the best result compared with other existing saliency-detection algorithms.

4.3 Evaluation on SOD dataset

We also compared our method with other state-of-the-art methods on SOD dataset [29]. This dataset contains many multiple object images with complicated backgrounds. The P-R curve and F -measure histogram are shown in Fig. 7. From Fig. 7, we can see that the proposed model is able to generate comparable saliency-detection results and even give a superior performance to other existing methods.

5 Conclusion and discussion

In this paper, we have proposed a novel scheme combining multiple low-level saliency priors and local similarity cue for efficient and accurate saliency detection. We first utilize multi-low-level prior maps and propagation mechanism based on Cellular Automata to achieve reliable saliency detection, then saliency refinement is utilized to further augment the saliency maps more smoothly and preserve the salient object outlines completely. We have tested the

designed model on three publicly available datasets, and experimental results show that the proposed method is effective and efficient.

Acknowledgements We would like to thank Prof. Hui Yu in the School of Creative Technologies, University of Portsmouth, for providing technical editing and proofreading of the manuscript.

This work was supported by National Natural Science Foundation of China (NSFC) (61601427, 61602229, 61771230); Royal Society – K. C. Wong International Fellow; Natural Science Foundation of Shandong Province (ZR2016FM40); Shandong Provincial Key Research and Development Program of China (NO. 2017CXGC0701); Fostering Project of Dominant Discipline and Talent Team of Shandong Province Higher Education Institutions.

References

1. Achanta R, Hemami S, Estrada F et al (2009) Frequency-tuned salient region detection. *IEEE Computer Vision and Pattern Recognition*:1597–1604
2. Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S (2012) SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans Pattern Anal Mach Intell* 34(11):2274–2282
3. Chang K-Y, Liu T-L, Chen H-T, Lai S-H (2011) Fusing generic objectness and visual saliency for salient object detection. *Proc IEEE Int Conf Comput Vis*:914–921
4. Cheng MM, Mitra NJ, Huang X et al (2015) Global contrast based salient region detection. *IEEE Trans Pattern Anal Mach Intell* 37(3):569–582
5. Cheng M, Warrell J, Lin W, Zheng S, Vineet V, Crook N (2013) Efficient salient region detection with soft image abstraction. *Proc IEEE Int Conf Comput Vis*:1529–1536
6. Cholakkal H, Rajan D, Johnson J (2015) Top-down saliency with locality-constrained contextual sparse coding. *BMVC*
7. Goferman S, Zelnik-Manor L, Tal A (2012) Context-aware saliency detection. *IEEE Trans Pattern Anal Mach Intell* 34(10):1915–1926
8. Harel J, Koch C, Perona P (2006) 'Graph-based visual Saliency', *Advances in Neural Information Processing Systems*, pp. 545–552
9. He S, Lau RWH, Yang Q (2015) Exemplar-Driven Top-Down Saliency Detection via Deep Association. *IJCV* 115(3):330–344
10. Hou X, Harel J, Koch C (2012) Image signature: Highlighting sparse salient regions. *IEEE Trans Pattern Anal Mach Intell* 34(1):194–201
11. Hou X, Zhang L (2007) Saliency Detection: A spectral residual approach. *IEEE Conference on Computer Vision and Pattern Recognition*:1–8
12. Huang F et al (2017) Salient object detection via multiple instance learning. *IEEE Trans Image Process* 26(4):1911–1922
13. Itti L, Koch C, Niebur E (1998) A model of saliency based visual attention for rapid scene analysis. *IEEE Trans Pattern Anal Mach Intell* 20(11):1254–1259
14. Jian M, Lam K-M, Dong J, Shen L (2015) Visual-patch-attention-aware Saliency Detection. *IEEE Transactions on Cybernetics* 45(8):1575–1586
15. Jian M, Qi Q, Dong J, Yin Y, Lam K-M (2018) Integrating QDWD with Pattern Distinctness and Local Contrast for Underwater Saliency Detection. *J Vis Commun Image Represent* 53:31–41
16. Jian M et al (2018) Saliency Detection Using Quaternionic Distance Based Weber Local Descriptor and Level Priors. *Multimed Tools Appl* 77(11):14343–14360
17. Jiang H, Wang J, Yuan Z, Liu T, Zheng N, Li S (2011) Automatic salient object segmentation based on context and shape prior. *Proc Brit Mach Vis Conf*:1–12
18. Jing H, He X, Han Q, El-Latif AAA, Niu X (2014) Saliency detection based on integrated features. *Neurocomputing* 129(10):114–121
19. Jung C, Kim C (2012) A unified spectral – domain approach for saliency detection and its application to automatic object segmentation. *TIP* 21(3):1272–1283
20. Kim J, Han D, Tai Y, Kim J (2014) Salient region detection via high-dimensional color transform. *IEEE Conference on Computer Vision and Pattern Recognition*:883–890
21. Li J, Chen H, Li G et al (2015) Salient object detection based on meanshift filtering and fusion of colour information. *IET Image Process* 9(11):977–985

22. Li J et al (2010) Probabilistic multi-task learning for visual saliency estimation in video. *Int J Comput Vis* 90(2):150–165
23. Liu X, Huet B (2016) Event-based cross media question answering. *Multimed Tools Appl* 75(3):1495–1508
24. Liu T, Sun J, Zheng N et al (2011) Learning to detect a salient object. *IEEE Trans Pattern Anal Mach Intell* 33(2):353–367
25. Liu X, Wang M, Yin B, Huet B, Li X (2015) Event-Based Media Enrichment Using an Adaptive Probabilistic Hypergraph Model. *IEEE T Cybernetics* 45(11):2461–2471
26. Ma YF, Zhang HJ (2003) Contrast-based image attention analysis by using fuzzy growing'. *Proceedings of the Eleventh ACM International Conference on Multimedia*, pp. 374–381
27. Mahadevan V, Vasconcelos N (2009) Saliency-based discriminant tracking. *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, IEEE
28. Mahadevan V, Vasconcelos N (2010) Spatiotemporal saliency in dynamic scenes. *IEEE Trans Pattern Anal Mach Intell* 32(1):171–177
29. Movahedi V, Elder JH (2010) Design and perceptual validation of performance measures for salient object segmentation. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*:49–56
30. MSRA (2019). http://research.microsoft.com/en-us/um/people/jiansun/SalientObject/salient_object.htm
31. Peng H et al (2017) Salient object detection via structured matrix decomposition. *IEEE Trans Pattern Anal Mach Intell* 39(4):818–832
32. Qin Y et al (2015) Saliency detection via cellular automata. *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, IEEE
33. Scharfenberger C, Wong A, Fergani K, Zelek JS, Clausi DA (2013) Statistical textural distinctiveness for salient region detection in natural images. *Proc IEEE Conf Comput Vis Pattern Recog*:979–986
34. Shen X, Wu Y (2012) A unified approach to salient object detection via low rank matrix recovery. *Proc IEEE Conf Comput Vis Pattern Recog*:2296–2303
35. Shi J et al (2016) Hierarchical image saliency detection on extended CSSD. *IEEE Trans Pattern Anal Mach Intell* 38(4):717–729
36. Song H, Liu Z, Xie Y et al (2016) RGBD Co-saliency Detection via Bagging-Based Clustering. *IEEE Signal Processing Letters* 23(12):1707–1711
37. Song J, Zhang H, Li X, Gao L, Wang M, Hong R (2018) Self-Supervised Video Hashing With Hierarchical Binary Auto-Encoder. *IEEE Trans Image Processing* 27(7):3210–3221
38. Toshev A, Shi J, Daniilidis K (2007) Image matching via saliency region correspondences. *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, IEEE
39. Wei Y, Wen F, Zhu W, Sun J (2012) Geodesic saliency using background priors. *Proc Eur Conf Comput Vis*: 29–42
40. Wolfram S (1983) Statistical mechanics of cellular automata. *Rev Mod Phys* 55(3):601
41. Yan Q, Xu L, Shi J, Jia J (2013) Hierarchical saliency detection. *Proc IEEE Conf Comput Vis Pattern Recog*:1155–1162
42. Yang J, Yang MH (2017) Top-down visual saliency via joint CRF and dictionary learning. *IEEE Transactions on PAMI* 39(3):576–588
43. Yang C, Zhang L, Lu H (2013) Graph-regularized saliency detection with convex-hull-based center prior. *IEEE Signal Processing Letters* 20(7):637–640
44. Yang C et al (2013) Saliency detection via graph-based manifold ranking. *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, IEEE
45. Zhai Y, Shah M (2006) Visual attention detection in video sequences using spatiotemporal cues. In: *Proc. 14th ACM Int. Conf. Multimedia*, pp 815–824
46. Zhang H, Kyaw Z, Chang SF, Chua TS (2017) Visual Translation Embedding Network for Visual Relation Detection. *CVPR*:3107–3115
47. Zhang J, Sclaroff S (2016) Exploiting Surroundedness for Saliency Detection: A Boolean Map Approach. *IEEE Trans Pattern Anal Mach Intell* 38(5):889–902
48. Zhu W, Liang S, Wei Y, Sun J (2014) Saliency optimization from robust background detection. *Proc IEEE Conf Comput Vis Pattern Recog*:2814–2821



Muwei Jian received the PhD degree from the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, in October 2014. He was a Lecturer with the Department of Computer Science and Technology, Ocean University of China, from 2015 to 2017. Currently, Dr. Jian is a Professor and Ph.D Supervisor at the School of Computer Science and Technology, Shandong University of Finance and Economics. His current research interests include human face recognition, image and video processing, machine learning and computer vision. Prof. Jian was actively involved in professional activities. He has been a member of the Program Committee and Special Session Chair of several international conferences, such as *SNPD 2007*, *ICIS 2008*, *APSIPA 2015*, *EEECS 2016*, *ICTAI2016*, *ICGIP 2016*, *ICTAI 2017* and *ICTAI 2018*. Dr. Jian has also served as a reviewer for several international SCI-indexed journals, including *IEEE Trans.*, *Pattern Recognition*, *Information Sciences*, *Computers in Industry*, *Machine Vision and Applications*, *Machine Learning and Cybernetics*, *The Imaging Science Journal*, and *Multimedia Tools and Applications*. Prof. Jian holds 3 granted national patents and has published over 40 papers in refereed international leading journals/conferences such as *IEEE Trans. on Cybernetics*, *IEEE Trans. on Circuits and Systems for Video Technology*, *Pattern Recognition*, *Information Sciences*, *Signal Processing*, *ISCAS*, *ICME* and *ICIP*.



Jing Wang is pursuing her Master's degree supervised by Prof. Muwei Jian, at the Department of Computer Science and Technology, Ocean University of China. Her research interests include image processing, pattern recognition, and underwater vision.



Junyu Dong received his B.Sc. and M.Sc. in Applied Mathematics from the Ocean University of China (formerly called Ocean University of Qingdao) in 1993 and 1999, respectively. He won the Overseas Research Scholarship and James Watt Scholarship for his PhD study in 2000 and was awarded a Ph.D. degree in Image Processing in 2003 from the School of Mathematical and Computer Sciences, Heriot-Watt University, UK.. Dr. Junyu Dong joined Ocean University of China in 2004. From 2004 to 2010, Dr. Junyu Dong was an associate professor at the Department of Computer Science and Technology. He became a Professor in 2010 and is currently the Head of the Department of Computer Science and Technology. Prof. Dong was actively involved in professional activities. He has been a member of the program committee of several international conferences, including the 4th International Workshop on Texture Analysis and Synthesis (associated with ICCV2005), the 2006 British Machine Vision Conference (BMVC 2006) and the 3rd International Conference on Appearance (Predicting Perceptions 2012). Currently, Prof. Dong is the Chairman of Qingdao Young Computer Science and Engineering Forum (YOCSEF Qingdao). He is a member of ACM and IEEE. Prof. Dong's research interest includes texture perception and analysis, 3D reconstruction, video analysis and underwater image processing.



Chaoran Cui received the B.S. degree in Software Engineering in 2010 and the Ph.D. degree in Computer Science and Technology in 2015, both from Shandong University, Jinan, China. During 2015–2016, he was a research fellow at Singapore Management University. He is now a Professor in School of Computer Science and Technology at Shandong University of Finance and Economics. His research interests include information retrieval, analysis and understanding on multimedia information, and computer vision.



Xiushan Nie received the Ph.D. degree from Shandong University, Shandong, China, in 2011. He is currently a Professor with Shandong University of Finance and Economics, Shandong, China. From 2013 to 2014, he was a Visiting Scholar with the University of Missouri-Columbia, Columbia, MO, USA. His research interests include multimedia retrieval and indexing, multimedia security, and computer vision.



Yilong Yin received the Ph.D. degree from Jilin University, Changchun, China, in 2000. From 2000 to 2002, he was a Post-Doctoral Fellow with the Department of Electronics Science and Engineering, Nanjing University, Nanjing, China. He is currently the Director of the data Mining, Machine Learning, and their Applications Group and a Professor of the School of Computer Science and Technology, Shandong University, Jinan, China. His research interests include machine learning, data mining, and computational medicine.