CrossMark

ORIGINAL ARTICLE

A novel optimization framework for salient object detection

Hanling Zhang · Min Xu · Liyuan Zhuo · Vincent Havyarimana

Abstract Visual saliency aims to locate the noticeable regions or objects in an image. In this paper, a coarse-to-fine measure is developed to model visual saliency. In the proposed approach, we firstly use the contrast and center bias to generate an initial prior map. Then, we weight the initial prior map with boundary contrast to obtain the coarse saliency map. Finally, a novel optimization framework that combines the coarse saliency map, the boundary contrast and the smoothness prior is introduced with the intention of refining the map. Experiments on three public datasets demonstrate the effectiveness of the proposed method.

Keywords Visual saliency · Boundary contrast · Smoothness prior · Optimization framework

1 Introduction

Visual saliency is an effective way to identify the most important and noticeable regions in a scene. In the last few decades, many researchers have devoted themselves to the study of visual attention [1,2] and many computational models have been developed. This new trend is motivated by the broad application of saliency detection in visual computer, such as image retrieval [3], image segmentation [4], object recog-

H. Zhang · M. Xu (✉) · L. Zhuo · V. Havyarimana
College of Information Science and Engineering, Hunan University, Changsha 410000, China
e-mail: xumin19891030@163.com

H. Zhang
e-mail: jt_hlzhang@hnu.edu.cn

L. Zhuo
e-mail: charleer@foxmail.com

V. Havyarimana
e-mail: havincent14@hnu.edu.cn

nition [5], object retarget [19] and adaptive compression of images [6]. Generally, there are three sub-fields (fixation prediction [7], salient object detection [28] and objectness proposals [29,30]) which can be considered as a part of visual saliency detection. Based on what the model is driven by, there are always two types of computational models. One is the bottom-up model, which is fast and data driven. The other is the top-down model that is always slower and task driven.

Recently, most of the works have taken much effort to build bottom-up saliency models on low-level image features. The most fundamental measure for visual saliency is the contrast computation. Depending on where the contrast is computed, previous methods can be categorized into local contrast [7–10] and global contrast [4,11–13].

The local methods investigate various contrast measures in a small local neighborhood of the pixel or region. Itti et al. [7] utilized the color, intensity and orientation image features to develop a multi-scale bottom-up saliency method, which is usually used for comparison and is a milestone in saliency detection. Harel et al. [8] introduced a method to non-linearly combine the local uniqueness maps from different feature channels to highlight conspicuity. Ma and Zhang [9] used an alternative local contrast analysis for saliency estimation. Moreover, Liu et al. [10] presented an algorithm which uses the multi-scale contrast in a difference-of-Gaussian pyramid. The methods based on local contrast can only achieve success in limited aspects. The edges of the salient objects are better than the object's interior, since the latter cannot be highlighted uniformly.

The global methods compute the pixel or region saliency at global scale with respect to the entire image. Zhai and Shah [11] calculated the color saliency with image histograms in the whole image region. Goferman et al. [4] considered four principles of human visual attention to exact the saliency map. Based on the global contrast, Cheng et al. [12] designed

Fig. 1 Visual examples of the proposed approach. **a** The input images. **b** The ground truth. **c** The saliency maps of our methods

(a) Input **(b)** GT **(c)** Ours
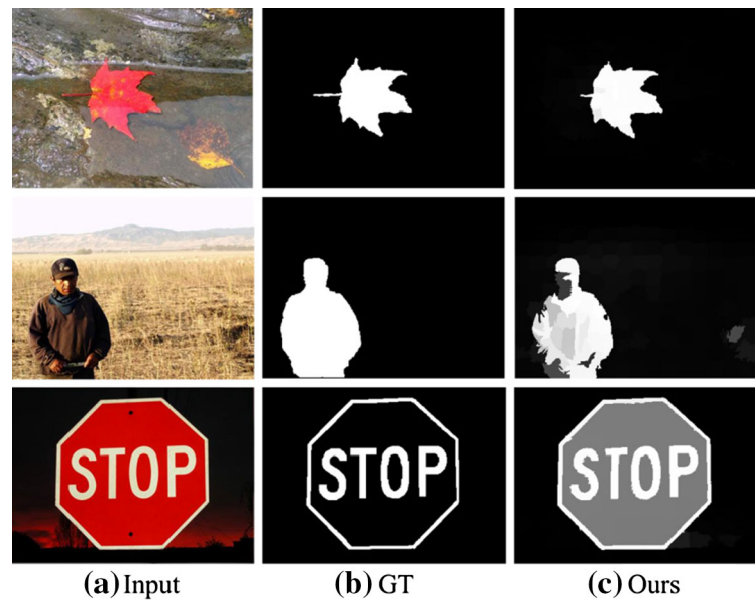
a saliency detection approach which involves either the color contrast or spatial coherence. Achanta et al. [13] proposed a frequency tuned method to generate the saliency map, with consideration of color difference between each pixel and the average value of the entire image in Lab color space. However, when the backgrounds are complex and the salient objects are small, there also exists difficulty in global method on distinguishing the salient object. Although global methods can alleviate the problem of highlighting the object uniformly which exists in local methods, these methods still have difficulties in highlighting the entire object uniformly.

Recently, a few methods which exploit the smoothness [14,15] item to refine the saliency quality have been proposed. Yang et al. [14] presented a novel bottom-up salient object detection approach by using contrast, center and smoothness priors. Based on quadratic programming framework, Li et al. [15] defined an approach that can adaptively optimize the regional saliency values on each specific image to simultaneously meet multiple saliency hypotheses on visual rarity, center bias and mutual correlation. However, these methods cannot make full use of the low-level information and the application of center bias still has some limitations.

To obtain a more robust result, and inspired by Yang et al. [14] and Li et al. [15], we propose a coarse-to-fine measure based on low-level information for defining image saliency. Firstly, we learn from [14] that the contrast and center priors were used to compute an initial prior map. Unlike most of the existing algorithms that refer to image center as priors, we estimate the center of the salient object by applying the convex hull of interest priors. Then we weight the initial prior map with boundary contrast to obtain the coarse saliency map. The boundary contrast is defined as the rarity of a region to boundary regions. Finally, we propose a novel optimiza-

tion framework that combines the coarse saliency map, the boundary contrast, and the smoothness prior to refine the map. This strategy can effectively suppress the background and uniformly highlight the salient object. We experimentally demonstrated that our method captures more the salient object than the state of the art methods [14,15] on famous benchmarks. Some visual saliency effects of the proposed method are shown in Fig. 1.

The rest of this paper is organized as follows. Details of the coarse saliency map are analyzed in Sect. 2, whereas the optimization framework is described in Sect. 3. The experiments on public datasets are preformed in Sect. 4 and the paper is concluded in Sect. 5.

2 The details of coarse saliency map

Taking the computational complexity into consideration, we over-segment the image into N super-pixels with the SLIC algorithm [16]. The measure can preserve the object boundaries better than the fixed size segmentation.

2.1 The initial map

Based on the color contrast and spatial coherence [14], the contrast prior map can be defined as a kind of regional rarity:

$$S_{\mathrm{co}}(i) = \sum_{j \neq i} \| c_i - c_j \| \cdot \exp\left(-\frac{\| p_i - p_j \|^2}{2\sigma_p^2} \right), \qquad (1)$$

where i and j denote the super-pixels, respectively, c_i and c_j are the mean color values of the corresponding super-pixel in CIE LAB color space, p_i and p_j are the average

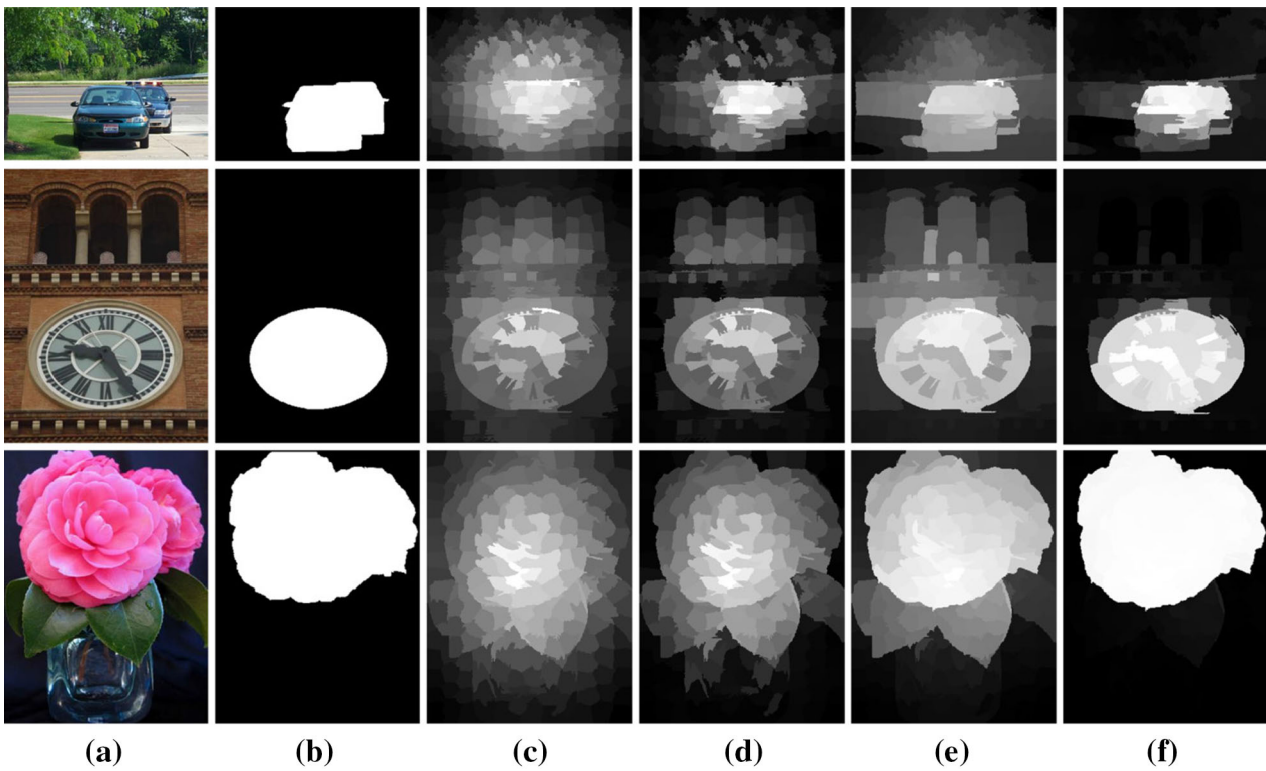| **(a)** | **(b)** | **(c)** | **(d)** | **(e)** | **(f)** |

Fig. 2 The comparison examples with Yang [14]. **a** The input images. **b** The ground truth. **c** The initial prior maps in [14]. **d** The coarse saliency maps by weighting the initial prior map with boundary contrast. **e** The final saliency maps (PBS) in [14]. **f** The refined saliency maps of the proposed approach

position whose values are normalized to [0, 1], and $\sigma_p^2 = 0.2$ indicates the strength of spatial coherence. Because of the absence of high-level priors, the initial map often incorrectly detects some background noises. Thus, we use the convex hull enclosing interesting points to estimate the general location of the salient object. Given the center (x_0, y_0), the convex hull-based center prior map can be defined as:

$$S_{ce}(i) = \exp\left(-\frac{\|x_i - x_0\|^2}{2\sigma_x^2} - \frac{\|y_i - y_0\|^2}{2\sigma_y^2}\right), \qquad (2)$$

where σ_x and σ_y control the horizontal and vertical variances, and we set $\sigma_x^2 = \sigma_y^2 = 0.15$ in our experiment. Then the initial prior map can be obtained by fusing the above two prior maps:

$$S_{in}(i) = S_{co}(i) \times S_{ce}(i). \qquad (3)$$

A visual saliency effect of the initial prior map is shown in Fig. 2c. By comparing the result, we note that the contrast prior map based on a super-pixel's contrast to all other super-pixels is inaccurate in many cases. Although the contrast prior map combines with the convex hull-based center prior map, there is still difficulty in suppressing the background efficiently.

2.2 The coarse saliency map weighted with boundary contrast

For most nature images, the background regions always appear smoothly and homogenously [18], while the salient pixels are usually grouped together [4]. From the photographic composition rules, we further observe that most photographers will not crop salient object along the view frame [18]. Thus, we can define the boundary contrast of a region as its color contrasts to the image boundary regions. It is close to 1 when the contrast is large and close to 0 when it is small. The definition is:

$$ctr_i = \exp\left(-\frac{1}{m}\sum_{j=1}^{m} w_{ij}\right), \qquad (4)$$

where m is the number of super-pixels on the image boundaries (we first use the SLIC algorithm to segment the image into N regions; then, we can obtain the boundary pixel set that was connected to the image boundary).

$w_{ij} = \exp(-\frac{\|c_i - c_j\|}{\sigma^2})$ is the color contrast ($\sigma^2 = 0.1$ empirically), whereas c_i and c_j are the mean color values of corresponding super-pixels and j represents the super-pixels in the image.
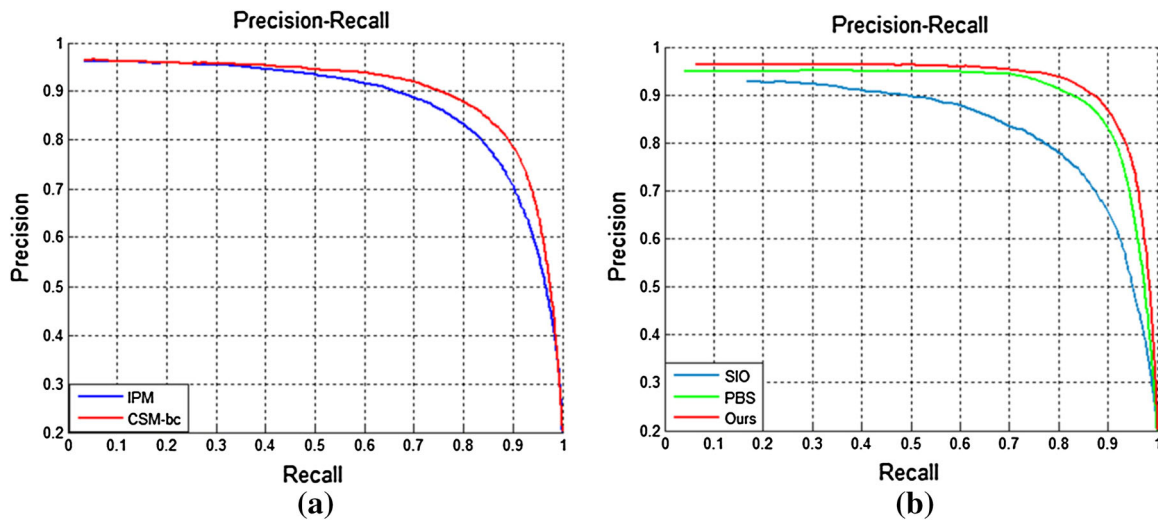
Fig. 3 Validation of the proposed approach. **a** Comparison of the proposed background contrast-weighted coarse saliency map (CSM_bc) with the initial prior map (IPM). **b** The performance estimation of the proposed optimization framework
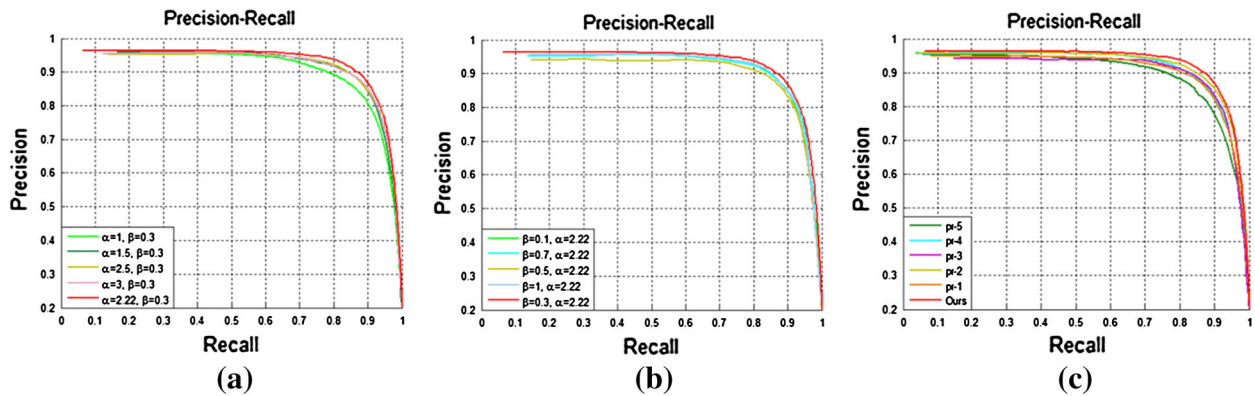

Fig. 4 The parameter setting and performance of different terms in the our optimization framework on the MSRA-1000 dataset. **a** The precision–recall comparison of different α on the refined saliency maps. **b** The precision–recall comparison of different β on the refined saliency maps. **c** The validation of different terms in our optimization framework

We weight the initial prior map with boundary contrast, which is defined as:

$$S_{\text{coar}}(i) = S_{\text{in}}(i) \cdot \text{ctr}_i. \tag{5}$$

According to Eq. (5), the object regions receive high ctr_i and the background regions receive small ctr_i, so the object regions are highlighted while the background regions are suppressed. This measure effectively enlarges the contrast between the object regions and background regions. Such improvement is clearly presented in Fig. 2d. The original initial prior map in [14] is messy when the background is complex (as shown in Fig. 2c). With the boundary contrast as weight, there is an obvious improvement.

The boundary contrast can suppress the background to a certain degree, while it is still bumpy and noisy. In the next section, we will propose a novel optimization framework to integrate these measures based on [15].

3 The optimization framework

From rarity hypothesis, center bias hypothesis and correlation hypothesis, Li et al. [15] transformed the problem of visual saliency estimation into an optimization framework. To combine multiple saliency cues or measures, we introduce a novel optimization framework that combines the coarse saliency map, the boundary contrast and smoothness prior to obtain the refined saliency map.

In this work, we model the saliency detection problem as the optimization of the saliency values of all the image super-pixels. The energy function is designed to assign the
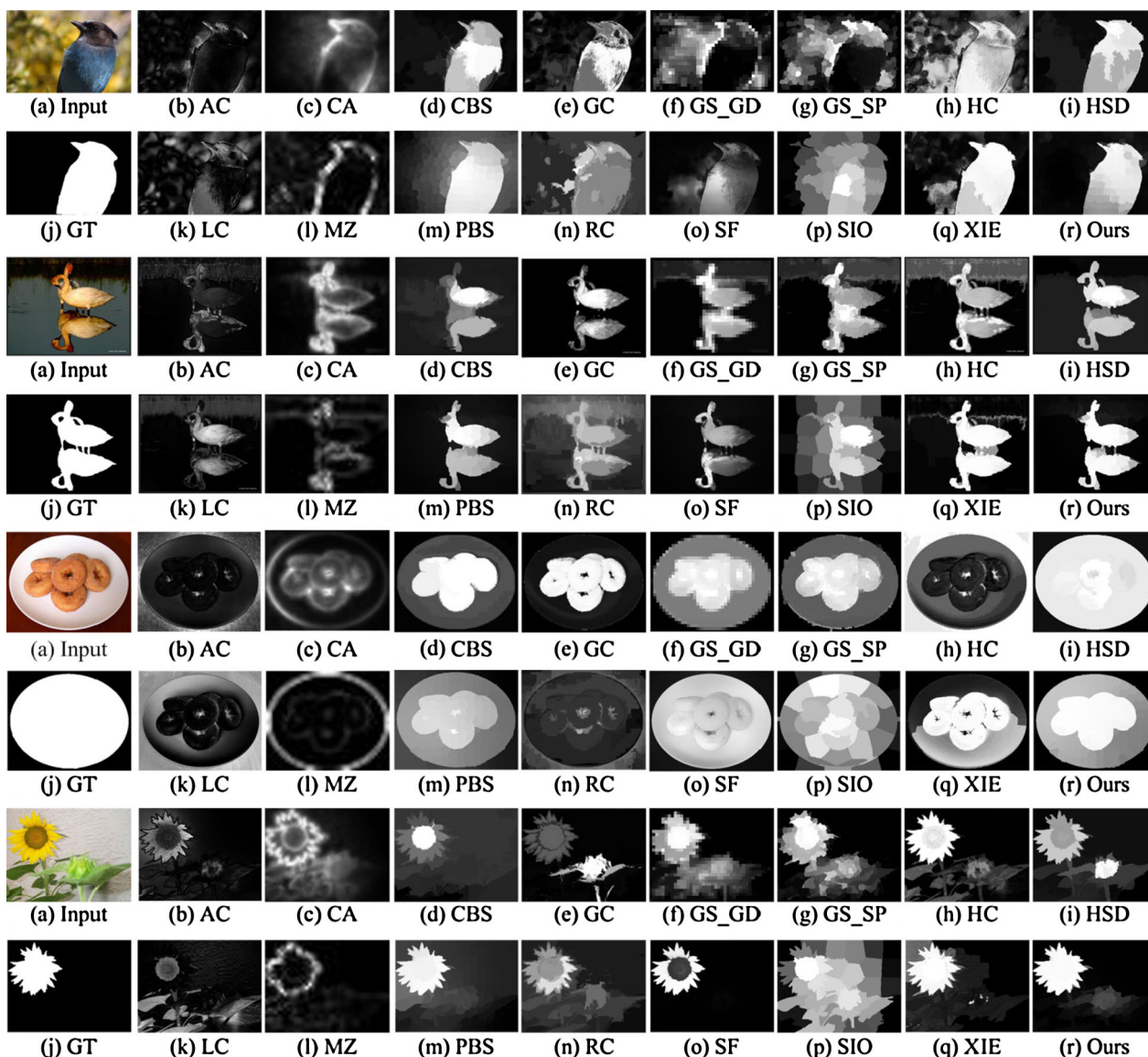
Fig. 5 Saliency detection results of different methods on the MSRA-1000 [13] dataset. The proposed approach consistently generates saliency maps close to the ground truth

salient super-pixel value 1 and value 0 to the background super-pixel. Let S_i be a saliency value of a super-pixel, then the energy function is defined as:

$$\arg\min_S \sum_{i=1}^{N} (1 - \text{ctr}_i) \cdot S_i^2 + \sum_{i=1}^{N} S_{\text{coar}}(i) \cdot (S_i - 1)^2$$
$$+ \sum_{i=1}^{N} \lambda_i \cdot (S_i - S_{\text{coar}}(i))^2 + \sum_{i=1}^{N} T_i \cdot (S_i - Z_i)^2,$$
$$+ \sum_{i,j=1}^{N} w_{ij}(S_i - S_j)^2. \qquad (6)$$

The five items are different constraints in the definition of saliency detection. According to the definition of ctr_i, it is close to 1 when the contrast is large and close to 0 when it is small. Thus, the value of $(1 - \text{ctr}_i)$ denotes the probability of super-pixel i which belongs to the background; it is large when the super-pixel belongs to the background and small when the super-pixel belongs to the salient object. The first item encourages a super-pixel to take a small value S_i with large background probability.

The following three items are all related to the coarse saliency map (Eq. 5). The second item indicates that a super-pixel with high value of $S_{\text{coar}}(i)$ takes a high value S_i (close to 1). The third item shows that the final saliency should not
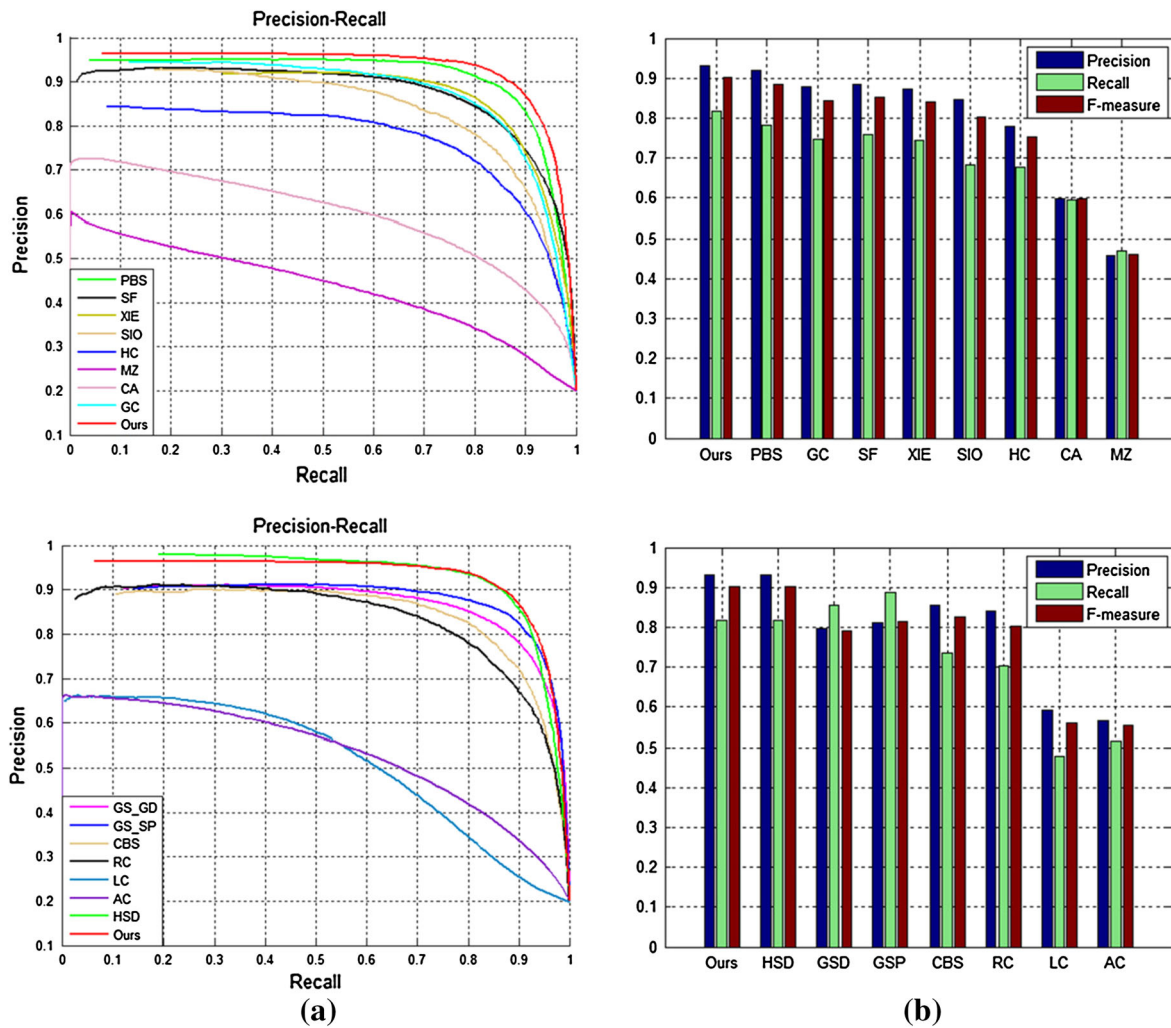
Fig. 6 Evaluation of the proposed work on the MSRA-1000 [13] dataset. Precision–recall curves in **a** is the comparison of different previous methods. **b** The precision, recall and F-measure

change too much from the coarse saliency map, whereas λ_i signifies that if a super-pixel's value is close to 1 or 0 in a coarse saliency map, it has a more significant impact on final saliency and it is defined as: $\lambda_i = \exp(-S_{\text{coar}}(i) \cdot (1 - S_{\text{coar}}(i))/\sigma_1^2)$.

The constraint parameter σ_1^2 is empirically set to 0.1. Inspired by [27] in the fourth item, T_i is the assuring constraint whose elements are 1 for certain pixels and 0 for all other pixels, and is given by:

$$T_i = \begin{cases} 1, & \text{if } (S_{\text{coar}}(i) \geq \alpha M) \\ 1, & \text{if } (S_{\text{coar}}(i) < \beta M), \\ 0, & \text{otherwise} \end{cases} \qquad (7)$$

where M represents the mean saliency value of the coarse saliency map. If $S_{\text{coar}}(i) \geq \alpha M$, we assume that the super-pixel i belongs to the foreground. Moreover, if $S_{\text{coar}}(i) < \beta M$, then i belongs to the background. Experimentally, we find that when we set $\alpha = 2.22$ and $\beta = 0.3$ (Fig. 4a, b),

the performance is stable. Z_i denotes the certain foreground pixels whose value is 1 and 0 for others. Inspired by closed-form solution [27], we admit that if the current super-pixel i belongs to a foreground region, the value of Z_i is 1, otherwise its value is 0. The last item encourages continuous saliency values. It indicates that a good saliency map should have similar saliency value between nearby super-pixels.

The minimum solution is computed by setting the derivative of the above energy function to zero. The five items can achieve impressive results and the optimization can be done fast due to the small number of super-pixels. Figure 2f shows the optimized results.

4 Experiments

We use the standard benchmark datasets: MSRA-1000 [13], SED1 [17] and BSD [18]. MSRA-1000 [13] is widely used
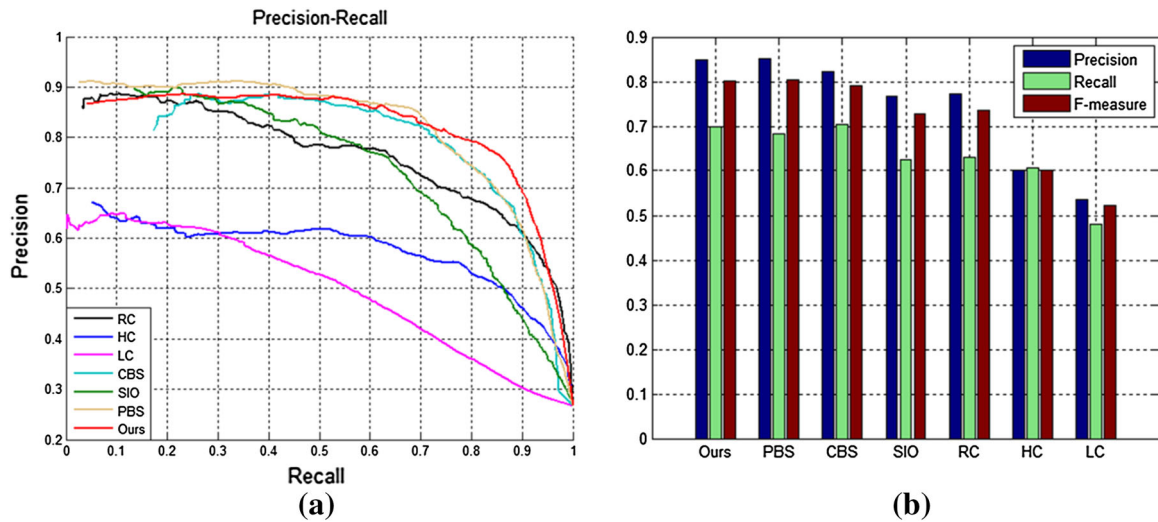
Fig. 7 Precision–recall curves and F-measure on SED1 [17] dataset to validate our algorithm. **a** Precision–recall curves of different methods. **b** Precision, recall and F-measure
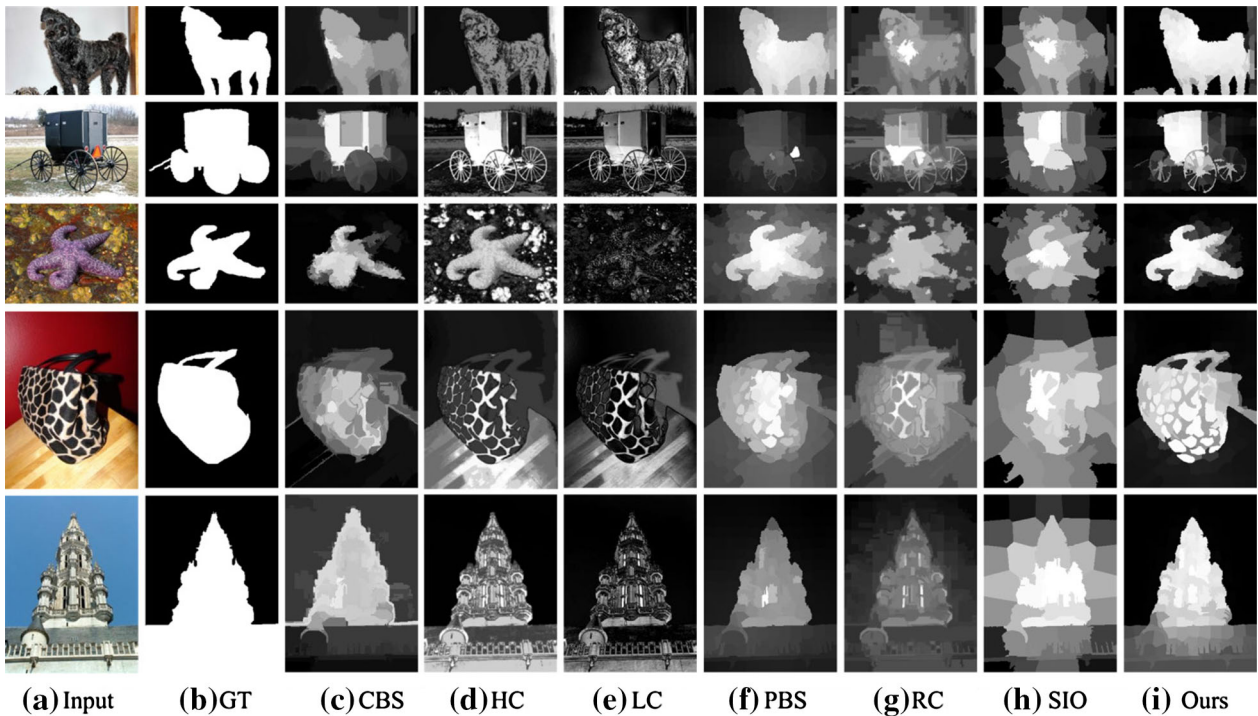

Fig. 8 Example results of different algorithms on SED1 [17] dataset to validate our approach

and relatively simple and contains 1,000 images with the corresponding accurate human-labeled binary masks for salient objects. The other two datasets are more challenging. The SED1 [17] contains 100 images and BSD [18] contains 300 images; these two datasets contain objects of different sizes and locations.

For performance evaluation, like many saliency detection models, we evaluate all methods through precision, recall and F-measure. Giving a saliency map with saliency value which is normalized to [0, 255], a set of binary images can be obtained by varying the threshold from 0 to 255. As a result, the precision–recall curve is generated based on the ground truth mask.

The F-measure is the overall performance of precision and recall, which can be measured as:

$$F_\beta = \frac{(1 + \beta^2) \; \text{Precision} \times \text{Recall}}{\beta^2 \; \text{Precision} + \text{Recall}}, \tag{8}$$

where $\beta^2 = 0.3$ according to [13].

(a) Input   **(b)** GT   **(c)** CBS   **(d)** HC   **(e)** LC   **(f)** GS_GD   **(g)** GS_SP   **(h)** PBS   **(i)** RC   **(j)** SIO   **(k)** Ours
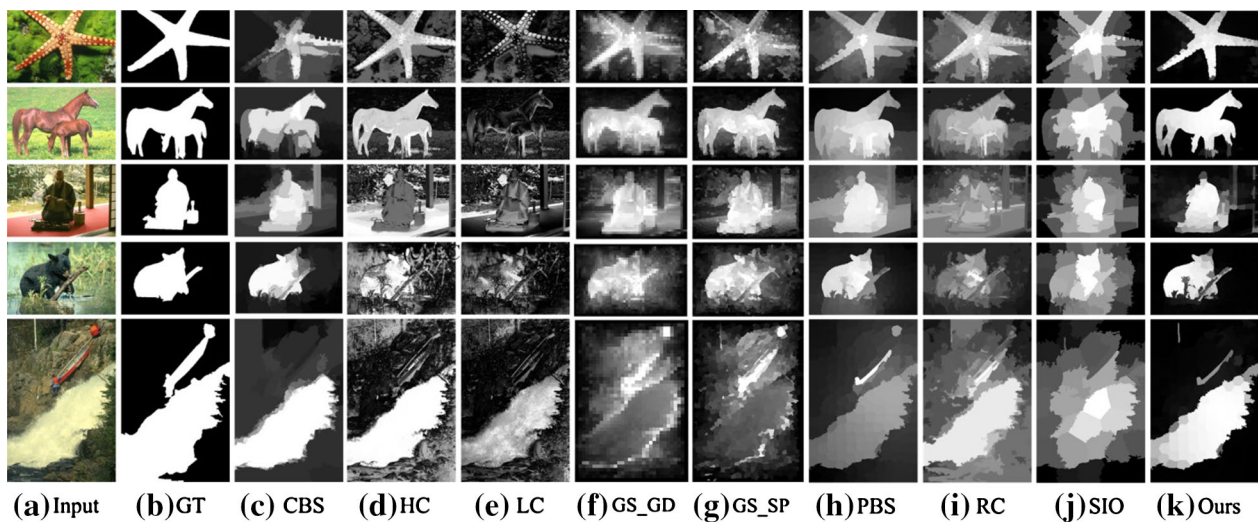
Fig. 9 Visual comparison of saliency maps on BSD [18] dataset. It can be observed that our methods can achieve better performance
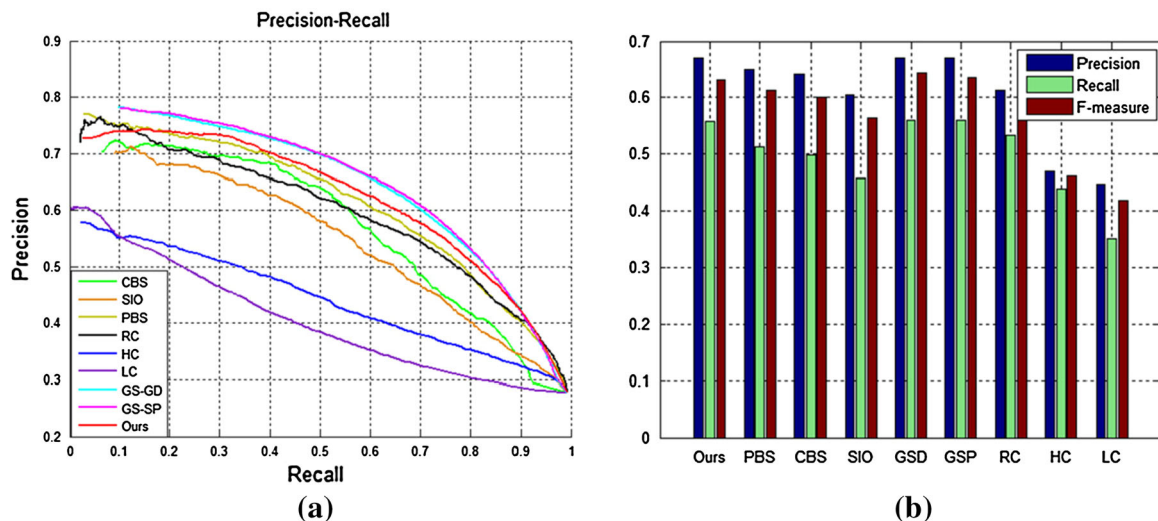

Fig. 10 Precision–recall curves and F-measure to measure the effectiveness of the proposed approach on BSD [18] dataset. **a** The comparison of different methods. **b** The evaluation of our method with F-measure

4.1 Validation of the proposed approach

To verify the effectiveness of the proposed approach, we compare the proposed background contrast-weighted coarse saliency map (CSM_bc) with the initial prior map (IPM) [14] by a precision–recall curve on MSRA-1000 [13] dataset. As shown in Fig. 3a, the red line represents the CSM_bc, which is higher than the blue line (IPM). Our coarse saliency maps have a better effect in precision and recall. This is because the boundary contrast effectively enlarges the difference between salient regions and backgrounds.

To estimate the performance of optimization framework, we compare our method with [14] and [15] with the precision–recall curve on MSRA-1000 [13] dataset firstly.

The results in Fig. 3b show that the proposed approach is significantly better than PBS [14] and SIO [15].

Then we analyze the effects of the different terms in Eq. (6) to validate the proposed optimization framework. We use quantitative result comparisons to analyze each term of the optimization framework. For example, we delete the first term and then present a precision–recall curve to show the effect. The precision–recall curves are shown in Fig. 4c. Because Eq. 6 has five terms, in Fig. 4c, the precision–recall curve shows pr–i when the ith item is deleted.

From the precision–recall curves in Fig. 4c, we can learn that our approach can achieve better performance. It is due to the full use of low-level information and our framework is more robust. In other words, the proposed method considers

more information of the image and presents the formulation in a more general way. The visual comparison results are shown in Fig. 5m, p, r.

4.2 Comparison with the other methods

We compare with the most recent 15 state-of-the-art methods on MSRA-1000 [13] dataset, including AC [20], RC [12], HC [12], CA [4], HSD [21], GC [22], MZ [23], SF [24], LC [11], XIE[25], GS_SD [18], GS_SP [18], CBS [26], SIO [15] and PBS [14]. To make a fair evaluation, we obtain the saliency maps of RC, HC, FT, LC, SR, AC, CA, GB, IT and MZ from [12]. For GS_SD, GS_SP and SF, we directly use the author-provided saliency results. For XIE, CBS, SIO as well as PBS, we run the authors' codes. The results of previous approaches and our algorithm are shown in Fig. 5. The precision–recall curve and F-measure are presented in Fig. 6a and b, respectively. From the results, we can see that our approach can achieve better performance, thanks to the effect of the optimization framework.

Based on the SED1 [17] dataset, we compare our method with six classic saliency models: RC [12], HC [12], LC [11], PBS [14], SIO [15] and CBS [26]. The precision–recall curve is shown in Fig. 7a, and the F-measure is shown in Fig. 7b. At the low recall values, the curve of PBS is slightly higher than our method. This is because we put more constraints in suppressing the background, so the precision values are a bit lower in the low recall values. We note that our method can suppress the background more effectively. Figure 8 indicates the visual results on the SED1 dataset.

Based on the BSD [18] dataset, we compare with eight previous models to estimate our methods, including RC [12], HC [12], LC [11], PBS [14], GS_GD [18], GS_GP [18], SIO [15] and CBS [26]. From Fig. 9, we can see the visual comparison with different algorithms. The comparison reveals that our method can achieve better performance. The precision–recall curve and F-measure are highlighted in Fig. 10. We note that most methods cannot obtain an appreciable result in precision and recall.

From the resulting curves, we note that the SED1 [17] and BSD [18] dataset are more challenging and need to be improved. The comparison results with other methods on three datasets indicate that our method significantly outperforms other classical methods in saliency detection.

However, like most methods, our method also contains some failure cases, e.g., when the salient object significantly touches the image boundary and there are complex backgrounds. Figure 11 presents the typical failure cases.

4.3 Computational efficiency

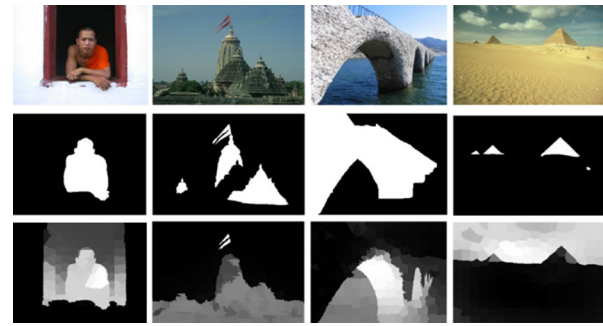We compare the performance of our method in terms of run time with several competitive accuracy methods or those sim-

Fig. 11 Typical failure cases of the proposed methods. *Top* the input images. *Middle* the ground truth. *Bottom* the saliency maps of the proposed method

Table 1 Comparison of average run time (seconds per image)

Method	Ours	PBS	SIO	CBS	XIE
Times (s)	0.855	0.749	0.843	2.03	140.679

ilar to ours on the MSRA-1000 dataset. The average run times of all the compared methods using a computer with Intel Pentium G630 2.70 GHz CPU and 2 GB RAM are presented in Table 1. Specifically, the super-pixel generation by the SLIC algorithm [16] spends 0.189 s, the coarse saliency map computation 0.644 s and the saliency map refining spends 0.022 s. The run time of the proposed method is a little slower than PBS [14] and SIO [15], due to the combination of more prior information in our model.

5 Conclusion

In this paper, we present a coarse-to-fine measure to model saliency. The boundary contrast is used to weight the initial prior map and obtain a more robust coarse saliency map. The optimization framework is applied to refine the coarse saliency map with the combination of coarse saliency map, the boundary contrast, and the smoothness prior. The experiment results on public datasets show that the proposed approach can effectively improve the results and achieve the start-of-the-art performance.

In future work, we will discuss the hierarchical measure which integrates more image features.

References

1. Treisman, A., Gelade, G.: A feature-integration theory of attention. Cognit. Psychol. **12**, 97–136 (1980)

2. Tsotsos, J.: What roles can attention play in recognition? In: Proceedings of Seventh IEEE International Conference Development and Learning, pp. 55–60 (2008)

3. Chen, T., Cheng, M., Tan, P., Shamir, A., Hu, S.: Sketch2photo: internet image montage. ACM Trans. on Graphics (2009)

4. Goferman, S., Zelnik-Manor, L., Tal, A.: Context-aware saliency detection. IEEE TPAMI, pp. 1915–1926 (2012)

5. Rutishauser, U., Walther, D., Koch, C., Perona, P.: Is bottom-up attention useful for object recognition? In: CVPR (2004)

6. Itti, L.: Automatic foveation for video compression using a neurobiological model of visual attention. IEEE Transactions on Image Processing, pp. 1304–1318 (2004)

7. Itti, L., Koch, C., Niebur, E.: A model of saliency based visual attention for rapid scene analysis. IEEE TPAMI, pp. 1254–1259 (1998)

8. Harel, J., Koch, C., Perona, P.: Graph based visual saliency. In: NIPS, pp. 545–552 (2006)

9. Ma, Y.F., Zhang, H.J.: Contrast based image attention analysis by using fuzzy growing. In: ACM Multimedia, pp. 374–381 (2003)

10. Liu, T., Yuan, Z., Sun, J., Wang, J., Zheng, N., X, T., H.Y., S.: Learning to detect a salient object. IEEE TPAMI **33**(2), 353–367 (2011)

11. Zhai, Y., Shah, M.: Visual attention detection in video sequences using spatiotemporal cues. In: ACM Multimedia, pp. 815–824 (2006)

12. Cheng, M.-M., Zhang, G.-X., Mitra, N.J., Huang, X., Hu, S.-M.: Global contrast based salient region detection. In: CVPR, pp. 409–416 (2011)

13. Achanta, R., Hemami, S.S., Estrada, F.J., Süsstrunk, S.: Frequency-tuned salient region detection. In: CVPR, pp. 1597–1604 (2009)

14. Yang, C., Zhang, L.H., Lu, H.C.: Graph-regularized saliency detection with convex-hull-based center prior. IEEE Signal Process. Lett. **20**(7), 637–640 (2013)

15. Li, J., Tian, Y., Duan, L.: Estimating visual saliency through single image optimization. IEEE Signal Process. Lett. **20**(9), 845–848 (2013)

16. Achanta, R., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: SLIC Superpixels, Tech. Rep. EPFL, Tech. Rep. 149300 (2010)

17. Alpert, S., Galun, M., Basri, R., Brandt, A.: Image segmentation by probabilistic bottom-up aggregation and cue integration. In: CVPR (2007)

18. Wei, Y.C., Wen, F., Zhu, W.J., Sun, J.: Geodesic saliency using background priors. In: ECCV (2012)

19. Wang, D., Li, G., Jia, W., Luo, X.: Saliency-driven scaling optimization for image retargeting. The Visual Computer, pp. 853–860 (2011)

20. Achanta, R., Estrada, F., Wils, P., et al.: Salient region detection and segmentation. In: ICVS (2008)

21. Yan, Q., Xu, L., Shi, J., Jia, J.: Hierarchical saliency detection. In: CVPR, pp. 1155–1162 (2013)

22. Cheng, M.-M., Warrell, J., Lin, W.-Y., Zheng, S., Vineet, V., Crook, N.: Efficient salient region detection with soft image abstraction. In: ICCV, pp. 1529–1536 (2013)

23. Ma, Y., Zhang, H.: Contrast-based image attention analysis by using fuzzy growing. ACM Multimedia (2003)

24. Perazzi, F., Krahenbuhl, P., Pritch, Y., et al.: Saliency filters: contrast based filtering for salient region detection. In: CVPR (2012)

25. Xie, Y.L., Lu, H.C., Yang, M.H.: Bayesian saliency via low and mid level cues. IEEE TIP **1**, 6 (2013)

26. Jiang, H., Wang, J., Yuan, Z., et al.: Automatic salient object segmentation based on context and shape prior. In: BMVC (2011)

27. Levin, A., Lischinski, D., Weiss, Y.: A closed-form solution to natural image matting. IEEE Trans. Pattern Anal. Mach. Intell. **30**(2), 228–242 (2008)

28. Alexe, B., Deselaers, T., Ferrari, V.: Measuring the objectness of image windows. IEEE TPAMI, pp. 2189–2202 (2012)

29. Borji, A., Sihite, D.N., Itti, L.: Salient object detection: a benchmark. In: ECCV (2012)

30. Cheng, M.-M., Zhang, Z., Lin, W.-Y., Torr, P.: BING: binarized normed gradients for objectness estimation at 300fps. In: IEEE CVPR (2014)

Hanling Zhang is an associated Professor at the Department of Information Sciences and Engineering at Hunan University since 2003. He has a PhD in Signal and Information Process from Northwestern Polytechnical University, China, 2003. From 2007 till 2008 he was a visiting professor at University of western Ontario, Canada. His research interests include machine learning, image processing.

Min Xu received the B.E. degree in communication engineering from Hunan Institute of Science and Technology, Yueyang, China, and he is an M.S. candidate student in Information science and Engineering at Hunan University, Changsha, China. His current research interests include saliency detection and feature extraction.

Liyuan Zhuo received the B.E. degree in communication engineering from Hunan International Economics University, Changsha, China, and she is an M.S. candidate student in Information science and Engineering at Hunan University, Changsha, China. Her current research interests include facial feature detection and objective tracking.

Vincent Havyarimana received his B.S. degree in Mathematics from University of Burundi, Bujumbura, in 2007 and the M.E. degree in Computer Science and Technology from Hunan University, Changsha, China, in 2011. Currently, he is a PhD candidate in the college of Information Science and Engineering, Hunan University, China. He is also a lecturer at one of the Universities of Burundi: "Ecole Normale Supérieure". His research interests include Wireless Communication and Mobile Computing.