IET Computer Vision

Research Article

ISSN 1751-9632
Received on 24th January 2018
Revised 28th August 2018
Accepted on 15th October 2018
E-First on 4th March 2019
doi: 10.1049/iet-cvi.2018.5013
www.ietdl.org

Salient object detection via reliable boundary seeds and saliency refinement

Xiyin Wu[1,2], Xiaodi Ma[1,2], Jinxia Zhang[3], Zhong Jin[1,2] ✉

[1]*School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, People's Republic of China*
[2]*Key Laboratory of Intelligent Perception and System for High-Dimensional Information of Ministry of Education, Nanjing University of Science and Technology, Nanjing 210094, People's Republic of China*
[3]*Key Laboratory of Measurement and Control of CSE, Ministry of Education, School of Automation, Southeast University, Nanjing 210096, People's Republic of China*
✉ *E-mail: zhongjin@njust.edu.com*

Abstract: Salient object detection can identify the most distinctive objects in a scene. In this study, a novel graph-based approach is proposed to detect a salient object via reliable boundary seeds and saliency refinement. A natural image is firstly mapped to a graph with superpixels as nodes. Saliency information is then diffused over the graph using seeds. For the reason that the boundary nodes may contain salient nodes, it is not appropriate to use all boundary nodes as the background seeds. Therefore, a boundary saliency measurement is proposed to obtain more accurate background seeds. After that, the information of background seeds is diffused by a two-stage scheme. A background-based map and a foreground-based map are generated based on the two-stage scheme. Furthermore, in order to enhance the detection accuracy, a refinement model is presented to fuse the information of background-based and foreground-based maps. Experiments on seven public datasets show the proposed algorithm out-performs the state-of-the-art salient object detection algorithms.

1 Introduction

Humans can find visually distinctive stimuli in a scene accurately and rapidly. This visual attention mechanism has been investigated by computer vision, neuroscience and cognitive psychology. The results of this mechanism can be depicted as a saliency map which expresses the priority of visual attention. Various computational models have been proposed to calculate such saliency map. These models can be applied to multiple computer vision/graphics applications as a first step: video compression [1], object recognition [2], image retrieval [3] and video retargeting [4].

The computational models can be divided into two branches. One branch is an eye fixation prediction [5]. It can predict points that people look at and its corresponding map includes several fixation dots. The other branch is salient object detection [6]. It pays attention to segment the most attractive objects in a scene and outputs a saliency map which can distinguish object boundaries. In the two branches, salient object detection can describe the whole outline of an object and has wider use than eye fixation prediction, which is the major concern in this paper.

Graph-based approaches become more and more popular to detect salient object for their simplicity and efficiency [7–9]. These approaches usually have three main factors: graph, seed and diffusion process. Firstly, a graph is used to describe an input image with the superpixels (or patches) as nodes. Secondly, the most representative nodes are selected as seeds. Finally, the seeds are used to diffuse their information to other nodes in the graph. For the reason that the effectiveness of diffusion result lies in the accuracy of the seeds, the seed selection method is important for graph-based approaches. Boundary prior is widely used to select seeds in recent works [8–10]. It derives from the basic principle of photographic composition, i.e. most cameramen not crop salient objects along the image frame [11]. As a result, regions near image boundaries are more likely to be background. However, as a salient object may appear near the border of an image, it is not reasonable to take all boundary nodes as background seeds. Otherwise, it may generate noisy background seeds and inaccurate diffusion result.

The diffusion process of the graph-based approach can be interpreted as various models, including conditional random fields (CRFs) [12, 13], quadratic energy model [14, 15], random walk model [9, 16], manifold ranking (MR) model [8, 17] and so on. In these models, the result of MR achieves good performance in terms of effectiveness and efficiency. In [8], Yang *et al.* propose a two-stage scheme based on MR to detect the salient object. In the first stage, four boundaries of an image are used as background seeds and diffused via the MR model. The diffusion result is binarised to acquire foreground seeds for the second stage. In this stage, another diffusion result is generated and converted to saliency map. The two diffusion results diffused in the two-stage scheme have complementary information to each other. However, saliency refinement using the two results is rarely discussed. Thus, the saliency map generated by [8] may mix with background noises, and cannot highlight the whole object uniformly at the same time.

In this paper, a graph-based approach is proposed to address the aforementioned problems. Fig. 1 shows the diagram of the proposed method. An input image is firstly segmented into superpixels as nodes. These nodes are linked by edges and a weight matrix is computed based on the distances of nodes. Then the measure of boundary saliency is used to pick out reliable boundary seeds. After that, the two-stage scheme is adopted. In the first stage, the saliency values are diffused from the side-specific boundary seeds to get four side-specific maps, which can form a background-based map. In the second stage, the background-based map is binarised by an adaptive threshold and then used to generate foreground seeds. A foreground-based map is obtained by diffusing the foreground seeds. Finally, the information of background-based map and the foreground-based map is integrated to get the refined map.

The main contributions of this paper are summarised as follows: (i) Salient regions in the boundary seeds are removed by using a novel boundary saliency measure. Thus, reliable boundary seeds can be obtained from the boundary saliency measure. (ii) The two diffusion results generated from the two-stage scheme are integrated by the proposed refinement model. This refinement model takes advantage of both the two diffusion results and achieves better results for salient object detection. (iii) The proposed approach is compared with fourteen state-of-the-art methods on seven datasets. The experimental results show the

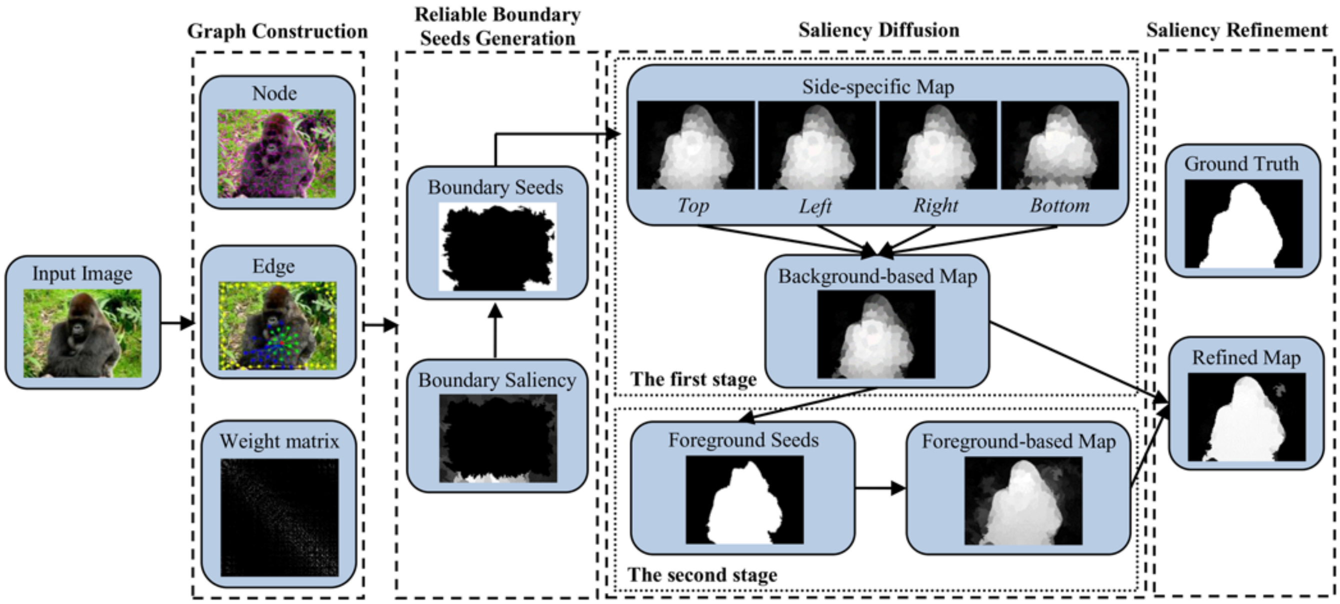

Fig. 1 *Diagram of the proposed approach. In the part of graph construction, the input image is segmented to superpixels as nodes (superpixel boundaries are in pink) and the edges of a node are illustrated. The given node (red) is connected to its spatial adjacent nodes (green) and colour adjacent nodes (blue). All boundary nodes (yellow) are connected with each other*

effectiveness of the proposed approach. The impact of each procedure and factor is analysed to obtain useful guidance.

The remaining of this paper is organised as follows. Related works are briefly reviewed in Section 2. In Section 3, the proposed approach is presented in detail. Section 4 gives experimental results and discusses the performance. The conclusions are given in Section 5.

2 Related work

The literature of visual attention is huge and can be categorised as eye fixation prediction and salient object detection. The former is surveyed in [18], and the latter is summarised in [19]. In this section, related work is reviewed belonging to the second category.

One of the earliest studies on salient object detection is proposed by Liu *et al.* [12], which treat visual attention as a binary segmentation issue. The frequency-tuned method [20] computes full resolution saliency maps using color and luminance features. Context-aware saliency [21] improves the result of respecting to their contexts in terms of spatial relationships and low-level feature. The unified low-rank matrix recovery model [22] decomposes the feature matrix of a picture into the low-rank part and sparse part, in which the sparse part represents the salient regions. The global contrast [23] assigns saliency value to each pixel using color statistics of the input image. Then an improved solution [24] is developed by integrating uniqueness and spatial distribution of image regions. The graph-based method [10] diffuses the information of salient seeds over the graph, and the diffusion result can be converted to a saliency map. Most recently, salient object detection takes some emerging cues into consideration, such as surroundness [25], depth [26] and objectness [27].

In the salient object detection models, some of them have been proposed based on graphs for the reason that a graph can reflect the relationships between image regions naturally. These models can be generally divided into four classes: CRFs, quadratic energy, random walk and MR. The typical model of CRFs uses an energy function with binary labels to embody unary and pairwise potentials. Liu and co-authors [12] propose a CRF model to combine three novel features: multiscale contrast, centre-surround histogram and colour spatial distribution. Mai *et al.* [13] extend CRF for saliency aggregation. In this way, the contribution of each saliency map can be modelled and the interaction between neighbouring pixels can be embodied. The binary label of CRFs has two limitations. The binary label makes CRFs inference complex. Besides, grade values are more suitable than the binary value to express saliency map. Therefore, the quadratic energy model is proposed to solve the two limitations. Chang *et al.* [14] construct a graph-based model to account for the relationship of objectness and saliency. The salient object detection result is improved by iteratively optimising an energy function. To ensure the information of seeds can be diffused to more distant areas and incorporated to external classifiers, Kong *et al.* [15] propose a model with quadratic Laplacian energy term. Saliency can be formulated as an alternative possibility of a random walk in the graph. Harel *et al.* [28] propose a random walk-based visual saliency method to identify salient regions via the frequency of node visits at equilibrium. Gopalakrishnan *et al.* [16] identify seed based on the hitting time from random walks to the most salient node, which is selected as the most isolated object in global. The goal of MR is to calculate a rank value for each node in the graph. The rank values of nodes are served as saliency values. MR has been applied to salient object detection in [8]. This method uses the two-stage scheme and achieves good performance so that a great many of the improvements are proposed. Li *et al.* [9] develop MR to regularised random walks ranking to calculate pixel-wised saliency maps. Tao *et al.* [17] propose an MR-based matrix factorisation method to embed seed labels in the coefficients. Inspired by the two-stage scheme, Gong *et al.* [10] employ a teaching-to-learn and learning-to-teach strategy to detect the salient object.

3 Salient object detection via reliable boundary seeds and saliency refinement

In this section, the proposed salient object detection approach is detailed. Our approach includes four major parts: the first one is graph construction, the second one focuses on reliable boundary seeds generation, the third one describes saliency diffusion process, and the last one refines the saliency map generated from the third parts.

3.1 Graph construction

An image I is divided into n superpixels by simple linear iterative clustering (SLIC) algorithm [29]. Then an undirected weighted graph $G = (\mathrm{V}, \mathrm{E})$ is constructed, where $\mathrm{V} = [v(i)]_{n \times 1}$ represents a set of nodes corresponding to superpixels and E represents the edges of the graph. The number of nodes n is set to 200 for all experiments.

The edges of the graph are added with three rules: (i) Same to [9], each pair of boundary nodes is enforced to connect. This rule

Table 1 Verification of two assumptions used in boundary saliency computation

Method	ASD	ECSSD	PASCAL-S	THUR15K	DUT-OMRON	DUTS	HKU-IS
images with salient boundary nodes	290	268	345	755	1510	1473	334
R_{obser1}	3.243	1.772	1.794	3.146	1.887	1.892	2.225
R_{obser2}	2.501	1.454	1.397	2.038	1.543	1.480	1.764

indicates the graph as the close-loop graph to decrease the geodesic distance of similar nodes. (ii) The immediate neighbour nodes and 2-hop neighbour nodes are connected. This rule exploits local spatial relationship since the neighbouring nodes share similar appearance with each other in general. (iii) For the reason that the second rule ignores the relationship among visually similar nodes, the adjacent nodes in colour space need to be connected. Let $v(i)$ and $v(j)$ are a pair of adjacent nodes if they meet the condition of $(D_c(i,j) + D_h(i,j))/2 < \tau$, where τ is a fixed threshold. As shown in Fig. 2, saliency maps are mixed with noises easily or cannot detect the object completely when τ is too small or too big. Therefore, τ is set to 0.15 in this paper. $D_c(i,j)$ is defined as the Euclidean distance $\| c(i) - c(j) \|$ between $v(i)$ and $v(j)$ in Lab colour space. $D_h(i,j)$ is the mean Chi-square distance χ^2 of three channels in the Lab colour space between two histograms $h(i)$ and $h(j)$ [10]

$$\begin{aligned} D_h(i,j) &= \frac{1}{3}\sum_{ch=1}^{3} \chi^2(h_{ch}(i), h_{ch}(j)) \\ &= \frac{1}{3}\sum_{ch=1}^{3}\sum_{k=1}^{K} \frac{(h_{ch,k}(i) - h_{ch,k}(j))^2}{h_{ch,k}(i) + h_{ch,k}(j)}, \end{aligned} \tag{1}$$

where ch is one of the channels in Lab colour space. $h_{ch,k}(i)$ denotes the kth histogram bin of $v(i)$ in the colour channel ch.

In the graph-based approaches, the similarity of two nodes is usually measured by a defined distance of the mean features of each region. According to the intuition that spatial nearby nodes or neighbouring regions in colour space are likely to have similar appearances. The weight of an edge is described by the distances of three kinds of features: colour, histogram and space. An entry of the weight matrix $\boldsymbol{W} = [w(i,j)]_{n\times n}$ is represented as follows:

$$w(i,j) = \exp\left(-\frac{\beta_c D_c(i,j) + \beta_h D_h(i,j) + \beta_s D_s(i,j)}{\delta^2}\right), \tag{2}$$

where δ is a constant controlling the strength of the weight. β_c, β_h and β_s are parameters to control the proportion of D_c, D_h and D_s, in which the colour distance D_c and the histogram distance D_h have been defined. D_s denotes the sine spatial distance proposed in [30]

$$D_s(i,j) = \sqrt{(\sin(\pi \cdot |x(i) - x(j)|))^2 + (\sin(\pi \cdot |y(i) - y(j)|))^2}, \tag{3}$$

where $x(i)$ and $y(i)$, respectively represent the horizontal and vertical coordinates of the node $v(i)$, which have been normalised to $[0, 1]$. The sine spatial distance reflects that the nodes belong to the opposite edges of the image is considered spatially close, while the nodes in the centre of the image are far from the boundary nodes.

3.2 Reliable boundary seeds generation

To all of the graph-based models, it is a challenging problem to determine which nodes should be selected as seeds, especially in the scenes with multiple objects. Recently, researchers move their attention to the background seeds rather than the salient seeds. In this case, the saliency map can be generated in an indirect way. That is, a background-based map is computed firstly, and then the saliency map can be generated by calculating the opposite value of the background-based map. The boundary prior states that objects seldom appear near the boundary of an image, most of the MR-related models use all of the four boundaries as background seeds

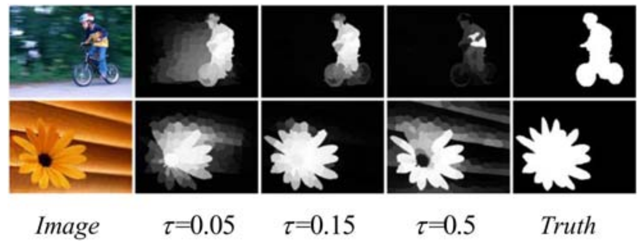

Fig. 2 *Visual comparisons of τ with different values. Image, $\tau = 0.05$, $\tau = 0.15$, $\tau = 0.5$, truth*

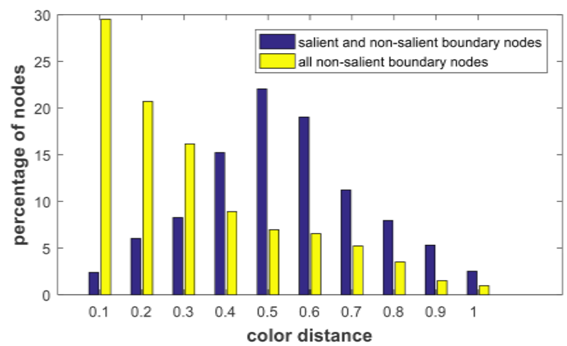

Fig. 3 *Colour distance distribution of boundary nodes*

[8, 10, 17]. However, salient objects may adjacent to image boundaries so that there are some noises in the background seeds.

In order to identify whether a boundary node belongs to the saliency nodes or not, the measure of boundary saliency is proposed according to two observations: (i) the salient object regions usually have high contrast to background regions [23]; and (ii) the salient object regions are much less connected to image boundaries than background regions [31]. Therefore, two corresponding assumptions are considered in the measure of boundary saliency. On the one hand, the colour and histogram distances among all boundary nodes can embody the first observation. To be specific, one of the boundary nodes probably belongs to salient object regions if it is far from other boundary nodes in the colour space. On the other hand, only a few nodes belong to the salient object may have large accumulated colour distances. To verify the two assumptions stated above, two corresponding experiments are tested on the images with salient boundary nodes from seven datasets. For the first assumption, the average colour distance between salient boundary nodes and other non-salient boundary nodes are compared with the average colour distance of all non-salient boundary nodes. The ratio of these two average distances is denoted as R_{obser1}. Moreover, the colour distance distribution of boundary nodes on all of the seven datasets is given (see Fig. 3) to verify the first assumption intuitively. For the second assumption, the average values of accumulated colour distances of salient boundary nodes and all boundary nodes are computed. The ratio of these two accumulated distances is denoted as R_{obser2}. As shown in Fig. 3, most of the colour distances between salient boundary nodes and other non-salient boundary nodes fall into $[0.4, 0.7]$, and most of the colour distances of all non-salient boundary nodes in the range of $[0.1, 0.3]$. From the experiment results presented in Table 1, R_{obser1} and R_{obser2} are all higher than 1, some of the ratio values are even higher than 2. In conclusion, most salient boundary nodes are far from non-salient nodes in colour space, and accumulated colour distances of salient boundary nodes are larger than non-salient nodes.

From the assumptions and the verification above, the boundary saliency S_B is defined as (see (4)) , where $b(i) \in \mathrm{B}$ is a boundary node. B denotes a boundary vector which is consisted of m boundary nodes. In (4), the average distances of colour and histogram features from $b(i)$ to other boundary nodes are calculated and then converted to a monotone increasing function by the log operation. This operation can enlarge the difference between the salient nodes and other nodes of the boundary, so that salient nodes of the boundary can emerge easily.

The boundary saliency provides how probable a boundary node belongs to the salient object. Therefore, when the boundary saliency of a node is bigger than a threshold, this node is served as a salient region. Let the vector $\boldsymbol{q}_B = [q_B(i)]_{n\times 1}$ denotes the reliable boundary seeds without salient noises. Then $q_B(i)$ can be defined as

$$q_B(i) = \begin{cases} 1, & S_B(b(i)) \le T_B, \\ 0, & \text{otherwise}, \end{cases} \tag{5}$$

where T_B is an adaptive threshold derived from [20]. Fig. 4 illustrates the effectiveness of using reliable boundary seeds. Figs. 4*b* and *f* show the examples of all boundary seeds and the reliable boundary seeds respectively, and Fig. 4*e* shows the boundary saliency. By using the boundary saliency, the salient noises near boundary are excluded. In this way, more reliable seeds can be provided to propagate in the diffusion step.

3.3 Saliency diffusion based on the MR model and two-stage scheme

The diffusion process with the MR model [32] is described as follows: Given an indication vector $\boldsymbol{q} = [q(i)]_{n\times 1}$, where $q(i) = 1$ if $v(i)$ is a seed. The remaining nodes are ranked with their relevance to the seeds. The goal of MR is to predict the value of unlabelled nodes by learning a diffusion function.

Let vector $\boldsymbol{f} = [f(i)]_{n\times 1}$ denotes the diffusion result, the optimal diffusion result of nodes is calculated by solving the following optimisation problem:

$$\begin{aligned} &\arg\min_f \frac{1}{2}\sum_{i,j=1}^{n} w(i,j)\| f(i) - f(j) \|^2 \\ &+\mu \sum_{i=1}^{n} d(i,i)\| f(i) - q(i)/d(i,i) \|^2, \end{aligned} \tag{6}$$

where $d(i,i)$ is an entry of diagonal degree matrix $\boldsymbol{D} = [d(i,i)]_{n\times n}$ with $d(i,i) = \sum_j w(i,j)$. The parameter μ controls the balance of the smoothness constraint (the first term) and the fitting constraint (the second term). We denote $\alpha = 1/(1+\mu)$. By setting the derivative of (6) to be zero, the optimised closed-form solution is represented as follows:

$$\boldsymbol{f} = (\boldsymbol{D} - \alpha \boldsymbol{W})^{-1}\boldsymbol{q}\,. \tag{7}$$

The two-stage scheme is adopted to diffuse saliency by using the MR model. In the first stage, a background-based map is integrated by four side-specific maps, i.e. the top, down, left and right boundary nodes are used to construct corresponding four maps. Different from [8], these four maps are computed using reliable boundary nodes rather than all boundary nodes as seeds. Taking top boundary as an example, the nodes located on the top boundary are tested by (5). Since the indication vector of the top boundary $\boldsymbol{q}_{BT}$ is generated, all other nodes obtain the diffusion result $\boldsymbol{f}_{BT}$ by (7). $\boldsymbol{f}_{BT}$ is normalised to $[0, 1]$ and its corresponding map can be written as

$$\boldsymbol{M}_T = \mathbf{1} - \bar{\boldsymbol{f}}_{BT}, \tag{8}$$

where $\mathbf{1} \in \mathbb{R}^{n\times 1}$ is a vector with all entries equal to 1, and $\bar{\boldsymbol{f}}_{BT}$ represents the normalised vector. The other three maps $\boldsymbol{M}_D$, $\boldsymbol{M}_L$ and $\boldsymbol{M}_R$ corresponding to the down, left and right boundary nodes are obtained in the same way. Then the four maps are integrated into the background-based map $\boldsymbol{M}_{BG}$:

$$\boldsymbol{M}_{BG} = \prod \boldsymbol{M}_i, \quad i \in \{T, D, L, R\} \tag{9}$$

In the second stage, foreground seeds are selected by using an adaptive threshold, which is set as the mean value of $\boldsymbol{M}_{BG}$. After that, the information of foreground seeds is diffused by (7) to obtain the diffusion result $\boldsymbol{f}_{FG}$. Same to the first stage, the values of $\boldsymbol{f}_{FG}$ are normalised to $[0, 1]$ to form the foreground-based map $\boldsymbol{M}_{FG}$:

$$\boldsymbol{M}_{FG} = \bar{\boldsymbol{f}}_{FG}, \tag{10}$$

where $\bar{\boldsymbol{f}}_{FG}$ represents the normalised vector.

In Fig. 4, the effectiveness of the two-stage scheme using reliable boundary seeds is illustrated. For the first stage, the background-based map diffused by all boundary seeds and reliable boundary seeds is compared in Figs. 4*c* and *g*. By using reliable boundary seeds, it is possible to highlight the salient nodes appearing on the boundaries in $\boldsymbol{M}_{BG}$. The nearby nodes of these salient nodes both in spatial and colour space can be highlighted during the diffusion process. Therefore, in the scene of the salient object adjacent to the boundary, the reliable boundary seeds can obtain more accurate diffusion results rather than all boundary seeds. The corresponding foreground-based maps of Figs. 4*c* and *g* are shown in Figs. 4*d* and *h*, respectively. The second stage improves the result of the first stage and more salient regions can be highlighted in this stage.

3.4 Saliency refinement

In the two-stage scheme, the background-based map $\boldsymbol{M}_{BG}$ and the foreground-based map $\boldsymbol{M}_{FG}$ are obtained via the reliable boundary seeds and foreground seeds. These two maps have complementary information to each other. To be specific, the background-based map highlights the salient objects, while the foreground-based map suppresses the background noises. Thus, $\boldsymbol{M}_{FG}$ can be refined by combining the information of these two maps.

$\boldsymbol{f}_{FG}$ is used to construct a new graph similar to the work of [33]. The definitions of node and edge are same as Section 3.1. The weight $\boldsymbol{W}_{FG} = [w_{FG}(i,j)]_{n\times n}$ of the new graph is evaluated by the Euclidean distance between $f_{FG}(i)$ and $f_{FG}(j)$:

$$w_{FG}(i,j) = \exp\left(-\frac{\| f_{FG}(i) - f_{FG}(j) \|}{\delta^2}\right). \tag{11}$$

Furthermore, two diagonal matrices are introduced to represent the information of $\boldsymbol{M}_{BG}$ and $\boldsymbol{M}_{FG}$. The information of $\boldsymbol{M}_{BG}$ is reflected by a diagonal matrix $\boldsymbol{D}_{BG} = [d_{BG}(i,i)]_{n\times n}$ with $d_{BG}(i,i) = \exp(-\gamma f_{BG}(i))$, where γ is a tuning parameter. Besides, the information of $\boldsymbol{M}_{FG}$ is represented by the diagonal matrix $\boldsymbol{D}_{FG} = [d_{FG}(i,i)]_{n\times n}$ with $d_{FG}(i,i) = \sum_j w_{FG}(i,j)$. A saliency refinement model is developed by these two diagonal matrices and the new graph. This model can generate the optimal refined result $\widehat{f}(i)$ by solving the following optimisation problem:

$$\begin{aligned} &\arg\min_{\widehat{f}} \frac{1}{2}\sum_{i,j=1}^{n} w_{FG}(i,j) \| \widehat{f}(i) - \widehat{f}(j) \|^2 \\ &+\mu \sum_{i=1}^{n} d_{FG}(i,i) \| \widehat{f}(i) - \frac{f_{FG}(i)}{d_{FG}(i,i)} \|^2 + \sum_{i=1}^{n} d_{BG}(i,i) \| \widehat{f}(i) \|^2, \end{aligned} \tag{12}$$

The three terms of the right-hand of (12) define costs from different constraints. The first one is a smoothness term to highlight the salient object uniformly. It denotes that nearby regions of a good saliency map should have similar values. Since $w_{FG}(i,j)$ computed by the value of $\boldsymbol{M}_{FG}$, the value of $\widehat{f}(i)$ maintains

$$S_B(b(i)) = -\log\left(1 - \frac{1}{2m}\sum_{j=1}^{m}(D_c(b(i), b(j)) + D_h(b(i), b(j)))\right), \tag{4}$$

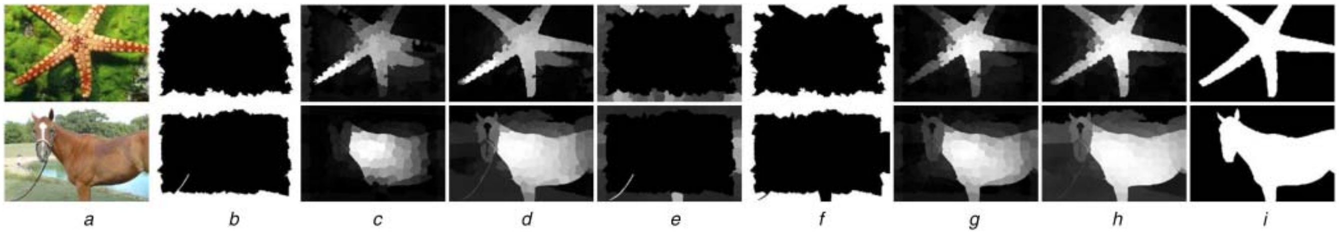

Fig. 4 *Effectiveness of using reliable boundary seeds and two-stage scheme*
(a) Input images, ***(b)*** Original boundary seeds, ***(c)*** Background-based map using all boundary seeds, ***(d)*** Foreground-based map using all boundary seeds, ***(e)*** Boundary saliency, ***(f)*** Reliable boundary seeds, ***(g)*** Background-based map using reliable boundary seeds, ***(h)*** Foreground-based map using reliable boundary seeds, ***(i)*** Ground truth

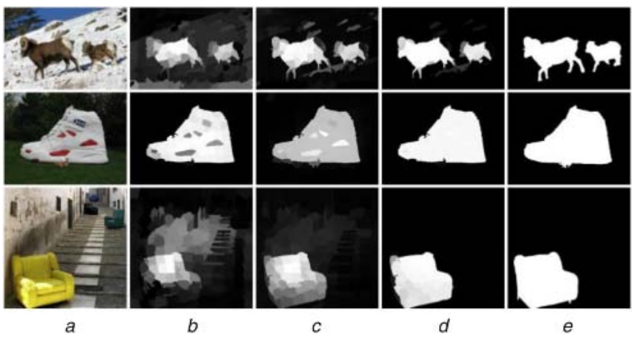

Fig. 5 *Benefit of the saliency refinement*
(a) Input images, ***(b)*** Background-based map, ***(c)*** Foreground-based map, ***(d)*** Saliency refinement, ***(e)*** Truth

the smoothness of $\boldsymbol{M}_{FG}$. The second one is a foreground term to encourage a node with small $f_{FG}(i)$ to have a low value $\widehat{f}(i)$ and vice versa. $f_{FG}(i)$ is the saliency value of node i in $\boldsymbol{M}_{FG}$. Therefore, the background nodes in $\boldsymbol{M}_{FG}$ should be suppressed and the foreground nodes in $\boldsymbol{M}_{FG}$ should be highlighted. This term denotes that $\widehat{f}(i)$ should not change too much from the foreground map value $f_{FG}(i)$. The third one is a background term which gives a confident soft value $\widehat{f}(i)$ with small $d_{BG}(i,i)$ (large $f_{BG}(i)$), whereas a low value $\widehat{f}(i)$ with large $d_{BG}(i,i)$ (small $f_{BG}(i)$). It indicates that a node belonged to the foreground in $\boldsymbol{M}_{BG}$ should have a confident value, and the background in $\boldsymbol{M}_{BG}$ should be suppressed in the refined result.

The minimum solution of (12) is computed by setting the derivate of this function to zero. The optimised solution is written as

$$\widehat{\boldsymbol{f}} = (\boldsymbol{D}_{FG} - \alpha \boldsymbol{W}_{FG} + \boldsymbol{D}_{BG})^{-1}\boldsymbol{f}_{FG}\,. \tag{13}$$

Thus, a refined map $\boldsymbol{M}$ can be obtained by normalising $\widehat{\boldsymbol{f}}$ to $[0, 1]$. By combining the information of background-based map and foreground-based map, the refined map has a better performance in highlighting the object and restraining the background, as shown in Fig. 5.

The main steps of our proposed algorithm are summarised in Algorithm 1.

Algorithm 1: Salient object detection via reliable boundary seeds and saliency refinement
Input: An image and required parameters.
Output: A saliency map $\boldsymbol{M}$.

1: Segment the input image into superpixels, construct a graph G with superpixels as nodes, add edges by three rules, and compute the weight of edges $\boldsymbol{W}$ by (2).

2: Compute the boundary saliency S_B by (4). S_B is binarised to generate the reliable boundary seeds vector $\boldsymbol{q}_B$ based on (5).

3: Diffuse saliency via the two-stage scheme. In the first stage, $\boldsymbol{q}_B$ are diffused to obtain the background-based map $\boldsymbol{M}_{BG}$ by (8) and (9). In the second stage, foreground seeds are selected by binarising $\boldsymbol{M}_{BG}$ with an adaptive threshold, and diffused to obtain the foreground-based map $\boldsymbol{M}_{FG}$ by (10).

4: $\boldsymbol{M}_{BG}$ and $\boldsymbol{M}_{FG}$ are integrated into get the refined result by (13). The refined result is normalised to get the final saliency map $\boldsymbol{M}$.

4 Experimental results

4.1 Experimental setup

4.1.1 Datasets: The proposed approach is evaluated on seven public datasets. The first one is ASD dataset [20] which contains 1000 pictures. This dataset is widely used for saliency detection. Most pictures in ASD dataset have only one salient object which usually has a strong contrast to the background. The second one is ECSSD dataset [34] which is an extend complex scene saliency dataset containing 1000 semantically meaningful but structurally complex images. The third one is PASCAL-S dataset [35] which contains 850 images with multiple objects. The fourth one is THUR15K dataset [36] which collects 15,000 pictures. Only 6233 pictures with pixel-wise masks are used in this paper. The fifth one is DUT-OMRON dataset [8]. It contains 5168 high- quality images which have one or more objects and relatively complex background. The sixth one and the last one are DUTS dataset [37] and HKU-IS dataset [38], respectively. These two datasets with testing and training sets can be used for supervised methods directly. The large-scale dataset DUTS contains 10,553 training images and 5019 testing images. The challenging HKU-IS dataset totally includes 4443 images, in which 2500 images are used for training, 500 images are used for validation and the remaining 1447 images are used for testing. To facilitate a fair comparison with other methods, only testing sets of DUTS dataset and HKU-IS dataset are utilised in the experiments.

4.1.2 Evaluation metrics: For a comprehensive evaluation, six metrics are used to evaluate the performance: precision–recall curve (PR curve), F-measure curve, receiver operating characteristic (ROC) curve, the mean of precision (MP) score, the mean of F-measure (MF) score and area under ROC curve (AUC) score. The saliency map is segmented with varying thresholds in the range $[0, 255]$, and compared the binary map with ground truth mask to compute the values of *PR curve*. In addition, the *F-measure curve* is used as an overall performance measurement calculated by a weighted average between precision (P) and recall (R):

$$F_\beta = \frac{(1+\beta^2)P \cdot R}{\beta^2 P + R} \tag{14}$$

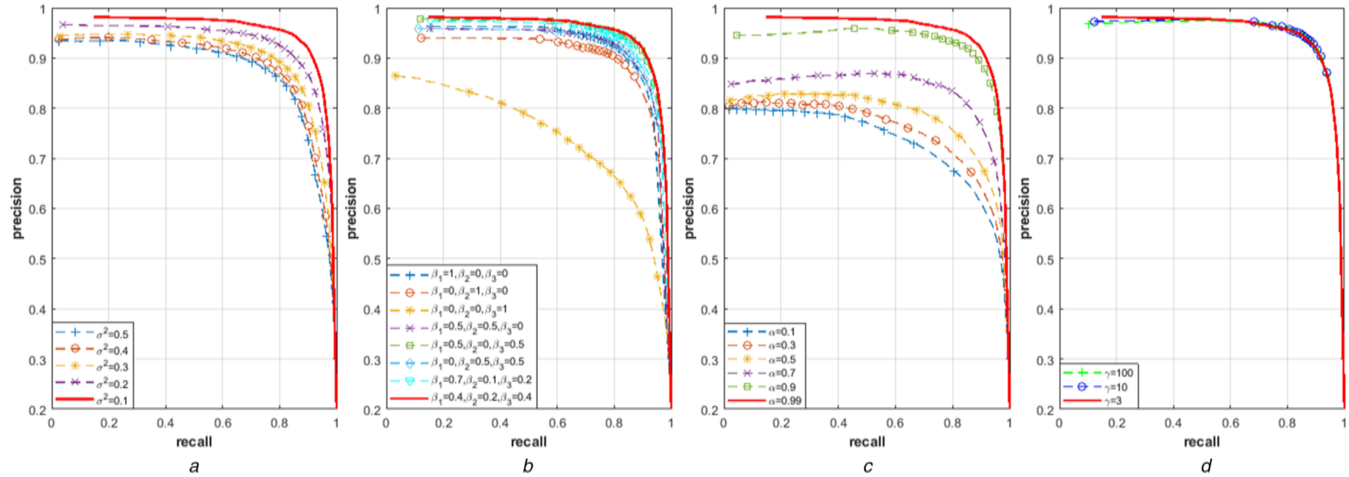

Fig. 6 *PR curves of different parameter setting*
(a) Parameter δ^2, ***(b)*** Parameters β_c, β_h and β_s, ***(c)*** Parameter α, ***(d)*** Parameter γ

where $\beta^2 = 0.3$ is used to emphasise the precision [19]. The ROC curve is plotted by the false positive rate (FPR) versus the true positive rate (TPR) varying thresholds in the range $[0, 255]$. While the three metrics details above are a two-dimensional representation of a method's performance, the other three metrics distil their information to a single scalar, naming MP, MF and AUC.

4.1.3 Parameter setting: The proposed approach has six main parameters. Different values of each parameter are tested on ASD with PR curve, and the results are shown in Fig. 6. In the part of graph construction, there are four parameters in (2): δ^2, β_c, β_h and β_s. The tuning parameter δ^2 is set to 0.1 based on Fig. 6*a*. Parameters β_c, β_h and β_s control the proportion of colour distance, histogram distance and spatial distance. They are set to 0.4, 0.2 and 0.4, respectively, according to Fig. 6*b*. In the part of saliency diffusion, the optimised closed-form solution of the MR model has a key parameter α, which is set to 0.99 based on Fig. 6*c*. It can emphasise the contribution of label consistency among neighbour nodes. The tuning parameter γ in the part of saliency refinement controls the importance of background-based map. The results are stable when γ varies in a large range. γ is set to 3 according to Fig. 6*d*.

4.2 Comparisons with the state-of-the-art methods

Fourteen state-of-the-art saliency detection approaches are evaluated for comparison, namely GB [28], GS [11], MR [8], RRW [9], TLLT [10], RBD [31], RC [23], SF [24], HC [23], LR [22], PCA [39], CA [40], WLR [41] and DSR [42]. The first six approaches are graph-based: GB and GS are classical methods; MR, RRW, TLLT and RBD are top-performance methods in recent years. The last eight are recent excellent approaches using variety models: RC is based on regional contrast, SF utilises contrast based filtering, HC is based on colour histogram, LR and WLR use low-rank matrix recovery theory, PCA uses principal component analysis, CA utilises cellular automata, and DSR is based on reconstruction errors. The source codes or executables for these approaches are available on the Internet.

4.2.1 Quantitative comparisons on classical datasets: In this section, the proposed method is validated on five classical datasets: ASD, ECSSD, PASCAL-S, THUR15K and DUT-OMRON. Fig. 7 shows the quantitative performance of various salient object detection methods evaluated by PR, F-measure and ROC curves. As depicted in Fig. 7*a*, the PR curves of the proposed method are higher than state-of-the-art methods on all of five datasets. In terms of the F-measure curves shown in Fig. 7*b*, the proposed method performs best on most of the thresholds when segmenting saliency maps, while it is lower than CA and WLR by varying thresholds in the range $[70, 220]$ on PASCAL-S dataset. As the ROC curves shown in Fig. 7*c*, our method gains relatively superior performance. In general, the proposed method is comparable to state-of-the-art methods evaluated by these three metrics.

The MP, MF and AUC of the proposed method and fourteen methods are listed in Table 2. The highest MP values achieved by our method are over 0.934, 0.818, 0.682, 0.600 and 0.593 on the five datasets. In terms of MF value, the proposed method obtains the best performance on the most of datasets except for PASCAL-S dataset. Although MR achieves the best MF on PASCAL-S dataset, it cannot work well on other datasets. GS, WLR and GB achieve the best AUC values on ASD, ECSSD and THUR15K datasets, respectively, while the proposed method provides the highest value of AUC on the rest of the three datasets. However, GS, WLR and GB cannot get good performance in terms of MP and MF because they usually just detect part of the object. Our method performs best on the DUT-OMRON datasets in term of MP, MF and AUC. Therefore, our method is validated strong potential in handling images with complex scenes.

4.2.2 Quantitative comparisons on challenging datasets: In this section, the proposed method is validated on two more challenging datasets: DUTS and HKU-IS. The quantitative comparisons in Fig. 8 are evaluated by PR, F-measure and ROC curves. As shown in Fig. 8*a*, the PR curve is the highest on HKU-IS dataset, while lower than DSR on DUTS dataset when the recall value is lower than 0.48. The F-measure curves in Fig. 8*b* show that our method has superior performance on the most of thresholds. Fig. 8*c* depicts the ROC curves in which our method is better than other methods. In total, our method also performs well in the challenging datasets.

4.2.3 Visual comparisons: Visual comparisons of some sample saliency maps are shown in Fig. 9. For images with simple scenes, some methods involve background regions in the saliency maps or miss some parts of the object, while the proposed method can pop out the entire salient object with uniform saliency values. For images with complex scenes, the proposed method can separate the foreground from the background successfully, while most methods cannot detect them. For the situation that the salient object near the boundary, our method can identify the whole regions of the salient object, especially the regions near the boundary. By contrast, other methods often fail in this situation. All of these results demonstrate the effectiveness of the proposed reliable boundary seeds method and the saliency refinement process.

4.2.4 Efficiency comparisons: The running time is tested on a machine with Intel(R) Xeon(R) E3–1231 v3 CPU @ 3.40 GHz and 8 GB RAM. The development environment is Matlab R2016a. The average runtime costs on ASD dataset are listed in Table 3. The proposed method takes about 0.317 s per image, which is

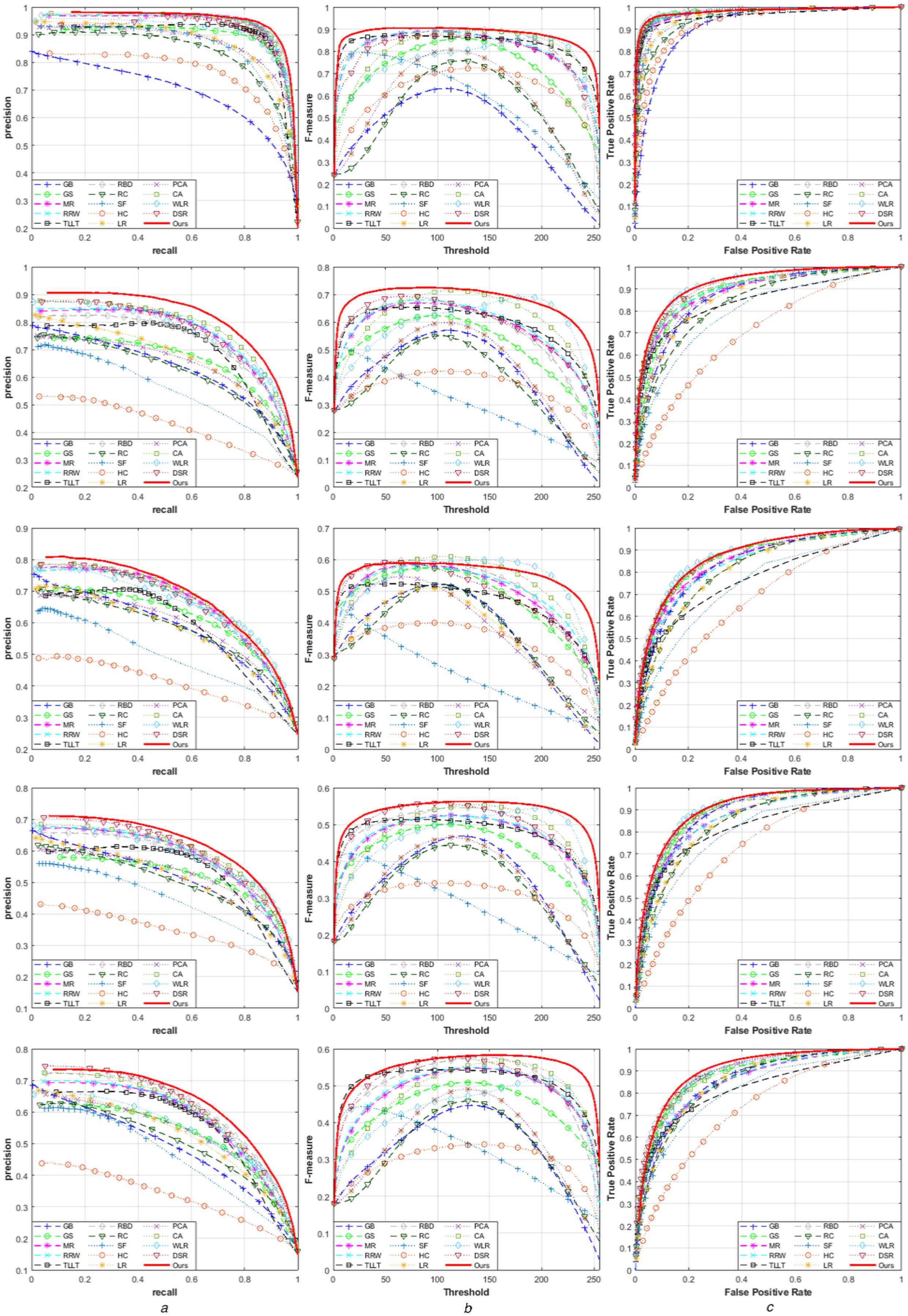

Fig. 7 *Quantitative comparisons on five datasets in terms of PR curve, F-measure curve and ROC curve. Top to down: ASD dataset, ECSSD dataset, PASCAL-S dataset, THUR15K dataset, DUT-OMRON dataset. Left to right:*
(a) PR curve, ***(b)*** F-measure curve, ***(c)*** ROC curve

Table 2 Quantitative comparisons on five datasets in terms of MP/MF/AUC

Method	ASD	ECSSD	PASCAL-S	THUR15K	DUT-OMRON
GB [28]	0.659/0.624/0.837	0.628/0.545/0.789	0.582/0.496/0.762	0.463/0.448/**0.828**	0.428/0.406/0.804
GS [11]	0.826/0.828/**0.877**	0.645/0.608/0.789	0.605/0.559/0.762	0.482/0.482/0.815	0.461/0.465/0.814
MR [8]	0.921/0.896/0.863	0.774/0.689/0.790	0.675/**0.602**/0.755	0.570/0.540/0.798	0.550/0.527/0.781
RRW [9]	0.920/0.897/0.866	0.771/0.697/0.796	0.676/0.587/0.747	0.574/0.543/0.796	0.550/0.529/0.782
TLLT [10]	0.913/0.881/0.845	0.770/0.666/0.744	0.679/0.533/0.671	0.584/0.526/0.742	**0.593**/0.546/0.741
RBD [31]	0.887/0.884/0.876	0.726/0.676/0.781	0.654/0.596/0.754	0.534/0.527/0.803	0.529/0.528/0.814
RC [23]	0.827/0.680/0.596	0.682/0.455/0.569	0.612/0.401/0.576	0.513/0.413/0.635	0.528/0.406/0.607
SF [24]	0.868/0.808/0.797	0.612/0.492/0.624	0.540/0.422/0.594	0.471/0.416/0.644	0.482/0.435/0.668
HC [23]	0.732/0.701/0.596	0.454/0.380/0.569	0.433/0.350/0.576	0.353/0.331/0.635	0.343/0.319/0.607
LR [22]	0.858/0.783/0.868	0.705/0.563/0.788	0.610/0.476/0.747	0.507/0.454/0.801	0.520/0.453/0.805
PCA [39]	0.821/0.797/0.876	0.662/0.578/0.791	0.601/0.525/0.751	0.498/0.481/0.821	0.480/0.462/**0.827**
CA [40]	0.885/0.876/0.875	0.750/0.702/0.815	0.663/0.596/**0.769**	0.546/0.535/0.822	0.516/0.509/0.808
WLR [41]	0.804/0.812/**0.877**	0.691/0.642/**0.819**	0.627/0.570/0.768	0.507/0.504/0.827	0.470/0.467/0.813
DSR [42]	0.874/0.856/0.871	0.750/0.690/0.785	0.655/0.584/0.743	0.554/0.540/0.803	0.535/0.524/0.802
ours	**0.934**/**0.912**/0.863	**0.818**/**0.734**/0.792	**0.682**/0.585/**0.769**	**0.600**/**0.570**/0.822	**0.593**/**0.566**/**0.827**

The best result for each dataset is shown in bold.

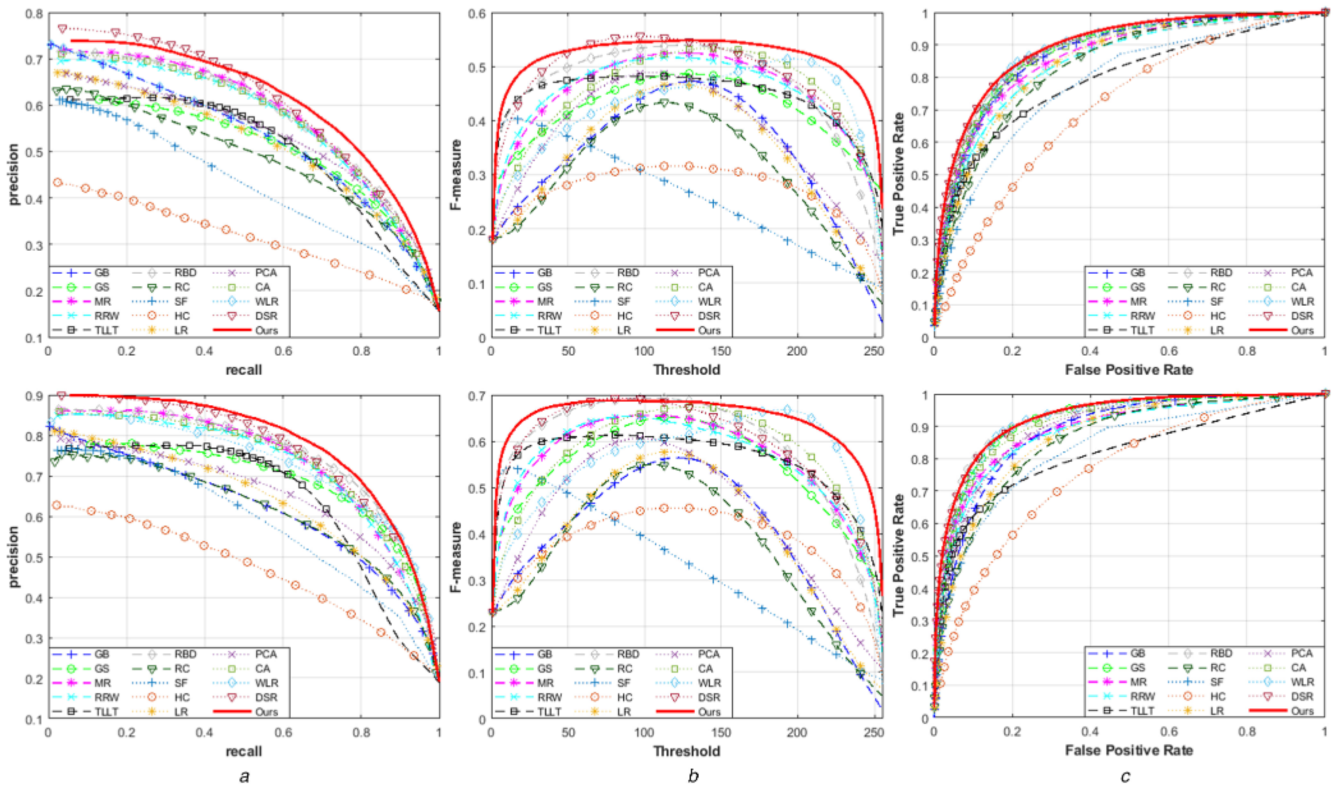

Fig. 8 *Quantitative comparisons on two challenging datasets in terms of PR curve, F-measure curve and ROC curve. Top to down: DUTS dataset, HKU-IS dataset. Left to right*
(a) PR curve, ***(b)*** F-measure curve, ***(c)*** ROC curve

significantly faster than GB, RRW, TLLT, LR, PCA, CA, WLR and DSR, although slower than GS, MR, RBD, RC, SF and HC. Considering overall evaluation performances, our method still outperforms other approaches.

4.3 Analysis of design options

To further understand the effects of the design options in the proposed approach, six experiments are examined on five datasets using MF value as the evaluation metric. As presented in Table 4, the six experiments are tested from two aspects: the results of three maps generated in our approach and the effectiveness of using the reliable boundary seeds. On the one hand, MF values of the background-based map $\boldsymbol{M}_{BG}$, the foreground-based map $\boldsymbol{M}_{FG}$ and the refined map $\boldsymbol{M}$ are growing on all of five datasets. Moreover, the performance of each map on ECSSD dataset is shown in terms of the PR curve (Fig. 10). The results reflect the effectiveness of the two-stage scheme and the saliency refinement. Since the proposed saliency refinement model combines the information of background-based map and foreground-based map, the model favourably improves the accuracy of the foreground-map. On the other hand, different maps diffused by all boundary seeds and the reliable boundary seeds are tested. Better results can achieve when the reliable boundary seeds are used as background seeds in the first stage of the two-stage scheme.

4.4 Limitation and discussion

Although the proposed method performs favourably for most of the cases, it may not work well if the colour of salient objects is very close to the background. Two failure cases are listed in Fig. 11. The reasons are concluded from two aspects. Firstly, the smoothness constraint of saliency diffusion model evaluates smoothness globally, it may miss the local information of the missing regions

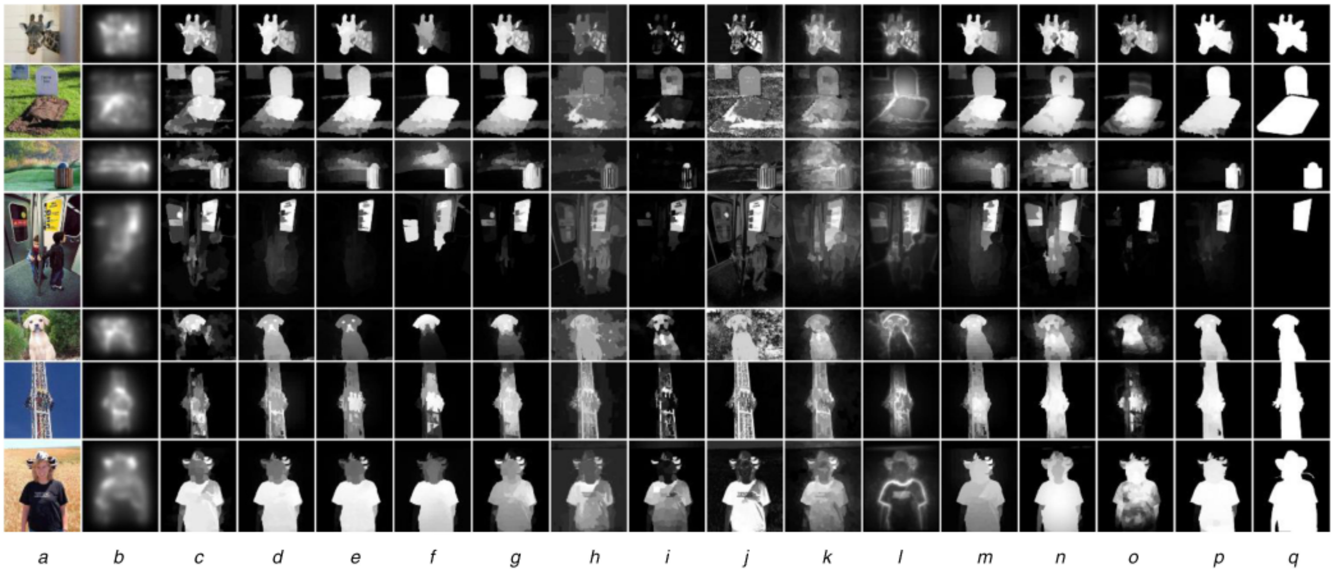

Fig. 9 *Visual comparisons of the proposed method versus fourteen state-of-the-art methods*
(a) Input image, ***(b)*** GB [28], ***(c)*** GS [11], ***(d)*** MR [8], ***(e)*** RRW [9], ***(f)*** TLLT [10], ***(g)*** RBD [31], ***(h)*** RC [23], ***(i)*** SF [24], ***(j)*** HC [23], ***(k)*** LR [22], ***(l)*** PCA [39], ***(m)*** CA [40], ***(n)*** WLR [41], ***(o)*** DSR [42], ***(p)*** Ours, ***(q)*** Truth

Table 3 Efficiency comparisons

Method	Ours	GB	GS	MR	RRW	TLLT	RBD	RC
time, s	0.317	0.659	0.234	0.311	0.906	1.127	0.258	0.135
code	Matlab	Matlab	Matlab	Matlab	Matlab	Matlab	Matlab	C++
method	Ours	SF	HC	LR	PCA	CA	WLR	DSR
time, s	0.317	0.145	0.016	11.862	2.354	0.567	1.060	2.143
code	Matlab	Matlab	C++	Matlab	Matlab	Matlab	Matlab	Matlab

Table 4 Examination of different design options on five datasets in terms of MF

Method	ASD	ECSSD	PASCAL-S	THUR15K	DUT-OMRON
M_{BG} with all boundary seeds	0.827	0.687	0.545	0.523	0.480
M_{FG} with all boundary seeds	0.890	0.707	0.556	0.553	0.535
M with all boundary seeds	0.911	0.731	0.584	0.564	0.563
M_{BG} with reliable boundary seeds	0.862	0.702	0.559	0.536	0.504
M_{FG} with reliable boundary seeds	0.907	0.729	0.574	0.568	0.543
M with reliable boundary seeds	0.912	0.734	0.585	0.570	0.566

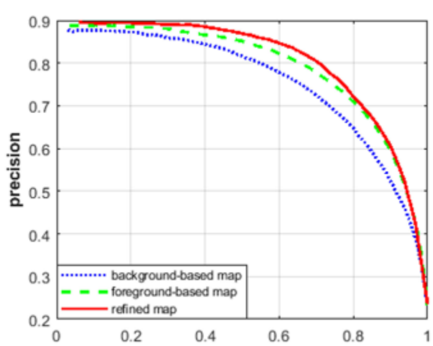

Fig. 10 *Performance of three maps generated in our method in terms of PR curve*

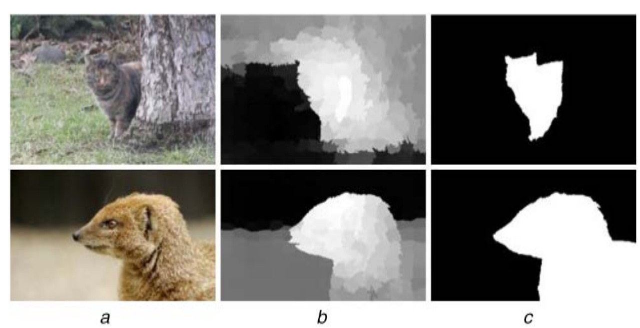

Fig. 11 *Failure cases of the proposed method*
(a) Input image, ***(b)*** Saliency map generated by the proposed method, ***(c)*** Ground truth

which have similar colour with the background. Besides, the computation of the weight matrix in the smoothness constraint mainly depends on the colour information. Therefore, it may be beneficial for the saliency diffusion model with another local smoothness constraint, and a weight matrix with more sophisticated information. Secondly, the boundary seeds generation way also relies on the information in colour space. So it fails to pick out accuracy seeds in this case. Therefore, other information is needed to add for comprehensive consideration.

Recently, deep learning methods develop rapidly in computer vision [43–45]. Although these methods achieve comparable results for saliency detection, they belong to supervised methods which depend on numerous training data. In addition, they are hard to learn visual information which is proved efficiently in hand-designed methods. Our method is unsupervised and no need for training data. It utilises graph to express the relationship between nodes and the underlying geometric structure. However, our method cannot perform satisfactorily as deep learning methods in some complex scenarios. To enhance the robustness of our method, the depth cue will be considered in the weight computation in the future.

5 Conclusion

In this paper, a graph-based approach is presented to detect a salient object. Instead of adopting all boundary nodes as background seeds, the boundary saliency measurement is proposed to exclude the foreground nodes in the two-stage scheme. Moreover, the two maps generated in the two-stage scheme are integrated by the saliency refinement model. The refined map can further highlight the salient object and suppress the background noises. Results of experiments on seven public datasets show that the proposed method outperforms the state-of-the-art saliency detection methods. In the future, more sophisticated information will be considered in the weight computation to improve the accuracy of our method. Moreover, the saliency diffusion model will be added a local smoothness constraint to distinguish salient objects from the similar background.

6 Acknowledgments

This work was partially supported by the National Natural Science Foundation of China under grant nos. 61872188, 61375007, 61233011, 91420201, 61472187, 61703100, National Basic Research Program of China under grant no. 2014CB349303, Natural Science Foundation of Jiangsu Province under grant no. BK20170692, and Postgraduate Research & Practice Innovation Program of Jiangsu Province under grant no. KYCX18_0426.

7 References

[1] Itti, L.: 'Automatic foveation for video compression using a neurobiological model of visual attention', *IEEE Trans. Image Process.*, 2004, **13**, (10), pp. 1304–1318

[2] Ren, Z., Gao, S., Chia, L.T., *et al.*: 'Region-based saliency detection and its application in object recognition', *IEEE Trans. Circuits Syst. Video Technol.*, 2014, **24**, (5), pp. 769–779

[3] Wang, P., Zhang, D., Wang, J., *et al.*: 'Color filter for image search'. Proc. ACM Int. Conf. Multi., Nara, Japan, 2012, pp. 1327–1328

[4] Rubinstein, M., Shamir, A., Avidan, S.: 'Improved seam carving for video retargeting', *ACM Trans. Graph.*, 2008, **27**, (3), pp. 443–453

[5] Li, J., Duan, L.Y., Chen, X., *et al.*: 'Finding the secret of image saliency in the frequency domain', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2015, **37**, (12), pp. 2428–2440

[6] Kim, J., Han, D., Tai, Y., *et al.*: 'Salient region detection via high-dimensional color transform'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Columbus, Ohio, USA., 2014, pp. 883–890

[7] Zhang, J., Ehinger, K.A., Wei H, , *et al.*: 'A novel graph-based optimization framework for salient object detection', *Pattern Recogn.*, 2017, **64**, pp. 39–50

[8] Yang, C., Zhang, L., Lu, H., *et al.*: 'Saliency detection via graph-based manifold ranking'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Portland, Oregon, USA., 2013, pp. 3166–3173

[9] Li, C., Yuan, Y., Cai, W., *et al.*: 'Robust saliency detection via regularized random walks ranking'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Boston, Massachusetts, USA., 2015, pp. 2710–2717

[10] Gong, C., Tao, D., Liu, W., *et al.*: 'Saliency propagation from simple to difficult'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Boston, Massachusetts, USA., 2015, pp. 2531–2539

[11] Wei, Y., Wen, F., Zhu, W., *et al.*: 'Geodesic saliency using background priors'. Proc. European Conf. Computer Vision, Firenze, Italy, 2012, pp. 29–42

[12] Liu, T., Yuan, Z., Sun, J., *et al.*: 'Learning to detect a salient object', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2011, **33**, (2), pp. 353–367

[13] Mai, L., Niu, Y., Liu, F.: 'Saliency aggregation: A data-driven approach'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Portland, Oregon, USA., 2013, pp. 1131–1138

[14] Chang, K.Y., Liu, T.L., Chen, H.T., *et al.*: 'Fusing generic objectness and visual saliency for salient object detection'. Proc. IEEE Int. Conf. on Computer Vision, Barcelona, Spain, 2011, pp. 914–921

[15] Kong, Y., Wang, L., Liu, X., *et al.*: 'Pattern mining saliency'. Proc. European Conf. on Computer Vision, Amsterdam, Netherlands, 2016, pp. 583–598

[16] Gopalakrishnan, V., Hu, Y., Rajan, D.: 'Random walks on graphs for salient object detection in images', *IEEE Trans. Image Process.*, 2010, **19**, (12), pp. 3232–3242

[17] Tao, D., Cheng, J., Song, M., *et al.*: 'Manifold ranking-based matrix factorization for saliency detection', *IEEE Trans. Neural Netw. Learn. Syst.*, 2016, **27**, (6), pp. 1122–1134

[18] Zhao, Q., Koch, C.: 'Learning saliency-based visual attention: a review', *Signal Process.*, 2013, **93**, (6), pp. 1401–1407

[19] Borji, A., Cheng, M.M., Jiang, H., *et al.*: 'Salient object detection: a survey', *IEEE Trans. Image Process.*, 2015, **24**, (12), pp. 5706–5722

[20] Achanta, R., Hemami, S., Estrada, F., *et al.*: 'Frequency-tuned salient region detection'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Miami, Florida, USA., 2009, pp. 1597–1604

[21] Goferman, S., Zelnik, M.L., Tal, A.: 'Context-aware saliency detection', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2012, **34**, (10), pp. 1915–1926

[22] Shen, X., Ying, W.: 'A unified approach to salient object detection via low rank matrix recovery'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Providence, Rhode Island, USA., 2012, pp. 853–860

[23] Cheng, M.M., Zhang, G., Mitra, N.J., *et al.*: 'Global contrast based salient region detection'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Colorado Springs, Colorado, USA., 2011, pp. 409–416

[24] Perazzi, F., Krahenbuhl, P., Pritch, Y., *et al.*: 'Saliency filters: contrast based filtering for salient region detection'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Providence, Rhode Island, USA., 2012, pp. 733–740

[25] Zhang, J., Stan, S.: 'Exploiting surroundedness for saliency detection: a Boolean map approach', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2016, **38**, (5), pp. 889–902

[26] Peng, H., Li, B., Xiong, W., *et al.*: 'RGBD salient object detection: a benchmark and algorithms'. Proc. European Conf. Computer Vision, Zurich, Switzerland, 2014, pp. 1–18

[27] Ye, L., Liu, Z., Li, L., *et al.*: 'Salient object segmentation via effective integration of saliency and objectness', *IEEE Trans. Multimedia*, 2017, **19**, (8), pp. 1742–1756

[28] Harel, J., Koch, C., Perona, P.: 'Graph-based visual saliency', Proc. Conf. Adv. Neural. Inf. Process, Vancouver, Canada, 2007, pp. 545–552

[29] Achanta, R., Shaji, A., Smith, K., *et al.*: 'SLIC superpixels compared to state-of-the-art superpixel methods', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2012, **34**, (11), pp. 2274–2282

[30] Zhang, J., Ehinger, K.A., Ding, J., *et al.*: 'A prior-based graph for salient object detection'. Proc. IEEE Int. Conf. on Image Processing, Paris, France, 2014, pp. 1175–1178

[31] Zhu, W., Liang, S., Wei, Y., *et al.*: 'Saliency optimization from robust background detection'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Columbus, Ohio, USA., 2014, pp. 2814–2821

[32] Zhou, D., Weston, J., Gretton, A., *et al.*: 'Ranking on data manifolds', Proc. Conf. Adv. Neural. Inf. Process, Vancouver, Canada, 2004, pp. 169–176

[33] Zhang, L., Yang, C., Lu, H., *et al.*: 'Ranking saliency', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2017, **39**, (9), pp. 1892–1904

[34] Shi, J., Yan, Q., Xu, L., *et al.*: 'Hierarchical image saliency detection on extended cssd', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2016, **38**, (4), pp. 717–729

[35] Li, Y., Hou, X., Koch, C., *et al.*: 'The secrets of salient object segmentation'. Proc. IEEE Int. Conf. Computer Vision, Columbus, Ohio, USA., 2014, pp. 280–287

[36] Cheng, M.M., Mitra, N.J., Huang, X., *et al.*: 'Salientshape: group saliency in image collections', *Visual Comp.*, 2014, **30**, (4), pp. 443–453

[37] Wang, L., Lu, H., Wang, Y., *et al.*: 'Learning to detect salient objects with image-level supervision'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Honolulu, Hawaii, USA., 2017, pp. 136–145

[38] Li, G., Yu, Y.: 'Visual saliency based on multiscale deep features'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Boston, Massachusetts, USA., 2015, pp. 5455–5463

[39] Margolin, R., Tal, A., Zelnik, M.L.: 'What makes a patch distinct?'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Portland, Oregon, USA., 2013, pp. 1139–1146

[40] Qin, Y., Lu, H., Xu, Y., *et al.*: 'Saliency detection via cellular automata'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Boston, Massachusetts, USA., 2015, pp. 110–119

[41] Tang, C., Wang, P., Zhang, C., *et al.*: 'Salient object detection via weighted low rank matrix recovery', *IEEE Signal Process. Lett.*, 2017, **24**, (4), pp. 490–494

[42] Lu, H., Li, X., Zhang, L., *et al.*: 'Dense and sparse reconstruction error based saliency descriptor', *IEEE Trans. Image Process.*, 2016, **25**, (4), pp. 1592–1603

[43] Zhao, R., Ouyang, W., Li, H., *et al.*: 'Saliency detection by multi-context deep learning'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Boston, Massachusetts, USA., 2015, pp. 1265–1274

[44] Wang, L., Lu, H., Ruan, X., *et al.*: 'Deep networks for saliency detection via local estimation and global search'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Boston, Massachusetts, USA., 2015, pp. 3183–3192

[45] Kuen, J., Wang, Z., Wang, G.: 'Recurrent attentional networks for saliency detection'. Proc. IEEE Conf. Computer Vision and Pattern Recognition, Las Vegas, Nevada, USA., 2016, pp. 3668–3677