Superpixel-Based Global Contrast Driven Saliency Detection in Low Contrast Images

Nan Mu[1] and Xin Xu[1,2(✉)]

[1] School of Computer Science and Technology,
Wuhan University of Science and Technology, Wuhan 430081, China
xuxin0336@163.com
[2] Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan University of Science and Technology, Wuhan 430081, China

Abstract. Due to the low signal to noise ratio, saliency detection in low contrast images has been a great challenge in computer vision. In this paper we propose a novel approach to detect salient object based on the computation of global saliencies in superpixel image blocks. This method tackles the image through a simple contrast measure, which first computes the global difference of two superpixels to obtain the resulting saliency map. Then, the map is refined by introducing the inter-superpixel similarity approach. The proposed model perfectly extracts the salient object in low contrast visibility conditions, which has been tested on three public datasets, as well as a nighttime image dataset. Experimental results demonstrate that the proposed method outperforms existing state-of-the-art saliency detection models.

Keywords: Superpixel segmentation · Low contrast · Salient object detection · Global saliency

1 Introduction

Visual saliency refers to a selection mechanism, the task of which is to extract the most important information for further processing. The research of saliency detection in natural images has proven to be useful for computer vision applications. With the development of various saliency models, it has witnessed tremendous advances in visual saliency detection in recent years. Most of these models focus on the contrast difference between salient objects and background region. Since the contrast between image elements (pixel, superpixel, or region) can be analogously used to compare the saliency of these elements, the natural images can be converted to saliency maps.

These models can work well in images with high contrast between foreground and background. But for detecting salient object in a relatively low contrast scene, they may face difficulties. Fig. 1 illustrates the saliency detection results using two state-of-the-art methods [1, 2]. The three testing images have low contrast between the visual salient objects and the background. In [1], Achanta *et al.* estimated the saliency by computing the difference between a pixel and the mean of the whole image in

H. Zha et al. (Eds.): CCCV 2015, Part I, CCIS 546, pp. 407–417, 2015.
DOI: 10.1007/978-3-662-48558-3_41

LAB color space. However, for the low contrast images, this approach fails to separate the salient object from the background, as in Fig. 1(b). In [2], Goferman *et al.* combined the local feature and global feature of the image patches to compute image saliency. This method can highlight the edges of the salient objects but miss the interior information, as in Fig. 1(c).

(a) Input (b) FT (c) CA (d) Proposed

Fig. 1. Examples of saliency detection results. (a) Input low contrast images with manually labeled rectangle. (b-c) Saliency maps obtained by two state-of-the-art methods [1, 2]. (d) Saliency maps obtained by the proposed method.

Generally speaking, most existing saliency detection approaches can be broadly classified into three main categories, in which the local contrast, global contrast, and the local-global contrast are considered, respectively.

1) The local contrast based saliency methods, which compute the distinctiveness of the image region in a local scope. The most popular method in this category is the saliency model proposed by Itti *et al.* [3], which computed the three local contrasts (luminance, color and direction) in different scales. Bruce and Tsotsos [4] utilized the information maximization approach to perform the local saliency computation. Han *et al.* [5] calculated the image saliency based on the sparse coding theory related to the local complexity.

2) The global contrast based saliency methods, which compute the distinctiveness of the image region over the entire image. Zhang *et al.* [6] measured the bottom-up saliency by extracting two features (difference of Gaussians and ICA-derived) of the entire image. Rosin [7] created the saliency map by utilizing the edge detection, distance transform, and thresholding approaches.

3) The local-global contrast based saliency methods, which consider both the local and global components, and then integrate them. Cheng *et al.* [8] proposed a global contrast model based on the region segmentation and refined the results by adopting a region based local contrast approach. Borji and Itti [9] proposed a saliency detection framework by measuring the local and global patch rarities and fusing them in a final map.

These contrast based methods have been successfully applied for proto-object detection. However, they perform poorly on low contrast images. It is a challenging task to acquire the effective features in low contrast images, and so far, only a handful of researches mention about the low contrast saliency detection. Wang *et al.* [10]

introduced a Salient Contrast Change (SCC) feature for object detection and tracking in low contrast videos. Han *et al.* [11] combined the multi-feature contrast weighted inhibition model and the fuzzy connection facilitation model to implement the contour detection in night vision images. Although these researches analyzed the influencing factors of night scene and put forward the effective measures to tackle with the night videos, there is still difficulty in extracting the salient objects accurately with low computational complexity.

To solve these problems, this paper presents a global contrast method based on local difference of each superpixel block, which can extract the salient object from low contrast images efficiently. The overview of the proposed method is shown in Fig.2. Unlike existing approaches, the proposed method does not require any training preprocessing, thus the computation is more efficient. This research utilizes the superpixel as the basic element, and operates on the simplified image by measuring the contrast difference between every two superpixel blocks in LAB color space. The proposed saliency model has a more preferable performance than the existing models for the saliency detection in low contrast images. The method has been tested on the MSRA dataset created by Liu *et al.* [12], the SED dataset created by Alpert *et al.* [13], the CSSD dataset created by Yan *et al.* [14], and the nighttime image dataset created by this project to corroborate its performance.

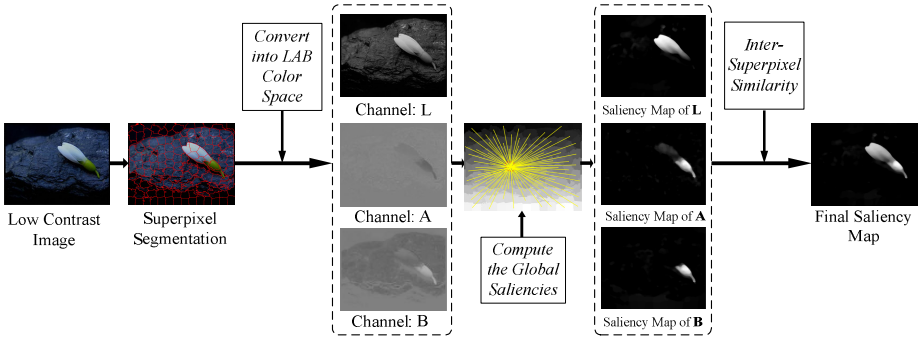


Fig. 2. An overview of the proposed framework.

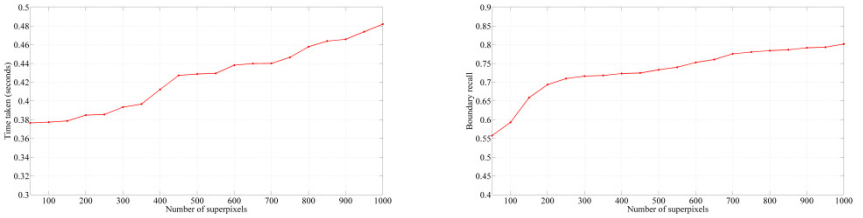
The rest of this paper is organized as follows. Section 2 describes the proposed salient object detection method. Section 3 presents the experimental comparison result of the proposed method with other existing saliency models. Finally, the conclusions are drawn in section 4.

2 Proposed Algorithm

The details of the proposed superpixel-based global contrast driven salient object detection algorithm are presented in this section.

2.1 Superpixel Segmentation

To simplify the operations, this study utilizes the superpixel segmentation method to partition the original image into a number of superpixels. We adopt the *simple linear iterative clustering* (SLIC) algorithm [15] to perform this process. The SLIC algorithm has an excellently perceptual characteristic, and the computational speed is very fast. In this work, we choose the optimal number (denoted as n) of the superpixels by analyzing the relation with the processing time and the boundary recall rate. Fig. 3 (a) and (b) plot the dependency of time taken and boundary recall rate on the number of superpixels, respectively.



(a) (b)

Fig. 3. The linear variation of (a) the time taken to generate superpixels and (b) the boundary recall rate influenced by the superpixel number.

From Fig. 3, which is tested on the mentioned publicly available datasets, we can observe that the time consumption of the SLIC algorithm is growing with the increasing of superpixel numbers, and the boundary recall rate also becomes higher. However, when the number of superpixels is larger than 200, the growing speed of recall rate will decrease. Thus, the optimal number n of superpixels is set to 200 in this study, this is sufficient for our work in learning the global difference between different regions and detecting the salient objects in low contrast images. It can not only guarantee a good boundary recall, but also shorten the computing time.

2.2 Global Contrast Approach

The original image is first converted into the CIELab space and decomposed into the respective L, A and B components. For each component, the superpixels are denoted as $SP_L(i)$, $SP_A(i)$, and $SP_B(i)$, respectively ($i = 1, \cdots, n$). The corresponding saliency value of each superpixel in every component is denoted as $SV_L(i)$, $SV_A(i)$, and $SV_B(i)$, respectively. We define the saliency value of superpixel $SP_L(i)$ by measuring the difference between each pixel value inside it (denoted as $SP_L(i; x, y)$) and the mean values of all other superpixels (denoted as $\overline{SP_L}(j), \quad j = 1, \cdots, n$), which is computed as:

$$SV_L(i) = \sum_{j=1}^{n} w(i, j) \cdot \left| SP_L(i; x, y) - \overline{SP_L}(j) \right|. \tag{1}$$

The weight $w(i, j)$ between superpixel $SP_L(i)$ and $SP_L(j)$ is obtained by computing the pixel number (denoted as $Num(j)$) in superpixel area $SP_L(j)$, and the Euclidean distance between the spatial center of $SP_L(i)$ and $SP_L(j)$, which are denoted as $c(i)$ and $c(j)$, respectively. The specific calculation is shown as follows:

$$w(i, j) = \frac{Num(j)}{|c(i) - c(j)|}. \tag{2}$$

The above algorithm is then executed in A and B components to compute the saliency value $SV_A(i)$ and $SV_B(i)$ of each superpixel. The saliency map (denoted as S_{Map}) is obtained by fusing the saliency map of L, A and B components via:

$$S_{Map} = \sqrt{SV_L \cdot SV_L + SV_A \cdot SV_A + SV_B \cdot SV_B}. \tag{3}$$

The resulting saliency map S_{Map} is normalized in the interval [0, 1], and the normalized feature map is calculated by:

$$S_{Map} = \frac{S_{Map} - \min(S_{Map})}{\max(S_{map}) - \min(S_{Map})}. \tag{4}$$

To further enhance the performance, the generated saliency map is smoothed by a median filter, which can better highlight the edges of the salient objects.

2.3 Internal Similarity Measure

We also introduce the inter-superpixel similarity measure [16] to refine the resulting saliency map. Each superpixel is assigned to a superpixel-level histogram $H_k(i)$, which is calculated based on the color quantization table with m entries. The histogram is normalized to have $\sum_{k=1}^{m} H_k(i) = 1$. The inter-superpixel similarity between two superpixel $SP(i)$ and $SP(j)$ is obtained by:

$$S(i, j) = \frac{S_{color}(i, j)}{|c(i) - c(j)|}. \tag{5}$$

The color similarity $S_{color}(i, j)$ is computed as the sum of intersection between each histogram:

$$S_{color}(i, j) = \sum_{k=1}^{m} \min\{H_k(i), H_k(j)\}. \tag{6}$$

The final saliency value for each superpixel is recalculated by exploiting the inter-superpixel similarity measure, so that the superpixels with higher similarity will have more similar values.

$$S'_{Map}(i) = \frac{\sum_{j=1}^{n} S(i,j) \cdot S_{Map}(j)}{\sum_{j=1}^{n} S(i,j)}. \tag{7}$$

The performance evaluation of the saliency maps obtained by the proposed method is described in the next section.

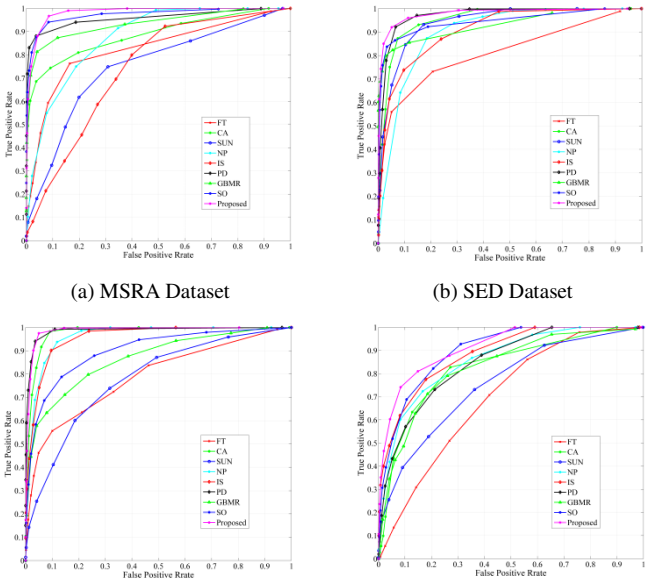
3 Experimental Results

A number of experiments were conducted to validate the performance of the proposed method on four datasets: (1) the MSRA dataset [12], in which the principle salient objects are labeled by different human subjects, (2) the SED dataset [13], which provides the ground truth, segmented by three human subjects. (3) the CSSD dataset [14], which is more challenging, including complex scenes, and (4) the nighttime image dataset created by the proposed research, which contains plenty of low contrast images in the evening, the resolution of these various images is $1280 \times 720 \times 24b$.

We compared our saliency model with eight existing state-of-the-art saliency models including *frequency-tuned* (FT) method [1], *context-aware* (CA) method [2], *saliency using natural statistics* (SUN) method [6], *non-parametric* (NP) method [17], *image signature* (IS) method [18], *patch distinction* (PD) method [19], *graph-based manifold ranking* (GBMR) method [20], and *saliency optimization* (SO) method [21].

In order to evaluate the performance of the proposed saliency model, we have introduced the *receiver operating characteristic* (ROC) graph to test the accuracy of generated saliency maps. The ROC graph is a two-dimensional graph which contains the *True Positive Rate* (TPR) and the *False Positive Rate* (FPR). The ROC curve is generated by plotting the obtained TPRs and FPRs, the ROC performance comparison of the eight methods and the proposed method is shown in Fig. 4, which are tested on the MSRA, SED, CSSD and the nighttime image dataset, respectively.

It can be seen from Fig. 4 that the proposed method has a better performance than other eight state-of-the-art saliency methods in MSRA, SED, CSSD, and nighttime image dataset, the overall performance will decline in the nighttime images which have a relatively low contrast. The *area under the curve* (AUC) is calculated to give an intuitive comparison. The AUC can indicate how well the generated saliency map predicts the human interesting area. Table 1 shows the AUC value of the various saliency models on the four datasets. It can be observed that the proposed model has state-of-the-art performance on the mentioned four datasets.



(a) MSRA Dataset (b) SED Dataset

(c) CSSD Dataset (d) Nighttime image Dataset

Fig. 4. The ROC performance plots for the four datasets.

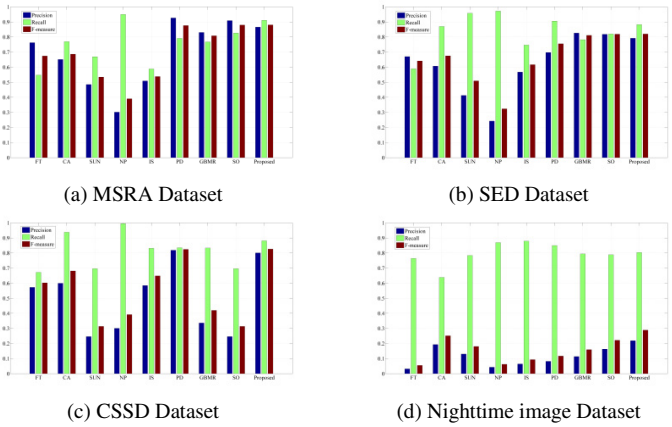
Table 1. The AUC performance of saliency maps from various saliency models on four datasets.

Datasets	Saliency models								
	FT	***CA***	***SUN***	***NP***	***IS***	***PD***	***GBMR***	***SO***	***Proposed***
MSRA	0.7515	0.9149	0.7188	0.8458	0.7396	0.9287	0.8722	0.9317	0.9551
SED	0.7326	0.9135	0.8806	0.8643	0.8356	0.9428	0.8469	0.9051	0.9458
CSSD	0.7382	0.9408	0.7280	0.9317	0.9365	0.9507	0.8039	0.8693	0.9518
Nighttime image	0.6978	0.7283	0.7533	0.8305	0.8506	0.8281	0.7991	0.8685	0.8767

For an objective comparison to quantitatively evaluate the performance for detecting the salient object, we introduce the precision, recall criteria, which calculated by comparing the binarized saliency map and the ground-truth mask. To further evaluate the accuracy of obtained binary mask of the saliency map, the F-*measure* is given by:

$$F_{measure} = \frac{(1+\beta^2)Precision \times Recall}{\beta^2 \times Precision + Recall}. \tag{8}$$

The proposed method uses $\beta^2 = 0.5$ to weigh the precision and recall. The comparison of precision, recall, and F-measure of these various methods are shown as:



(a) MSRA Dataset (b) SED Dataset

(c) CSSD Dataset (d) Nighttime image Dataset

Fig. 5. The Precision, recall, and F-measure performance comparison of various saliency models on four datasets.

As shown in Fig. 5, the F-measure value of the proposed method is relatively higher than the other eight methods, which indicates an excellent performance to predict the human eye gaze. The recall rate of the various saliency models is not high on the nighttime image dataset, the possible cause is that the salient objects in our dataset are too small, which results in a low F-measure performance.

The run-time performance is also considered to evaluate the efficiency of various algorithms. The experiment is measured on Intel 2.9GHZ CPU machine with 4GB RAM. All approaches use Matlab implementations. It can be observed from Table 2 that the run-time of IS method is time-saving, but can only generate the low resolution saliency maps. The computational complexity of the proposed method is slightly higher than the superpixel-based method GBMR, whereas our method can get more accurate estimations.

Table 2. The computational run-time (in second) of various saliency models on four datasets**.**

Datasets	Saliency models								
	FT	***CA***	***SUN***	***NP***	***IS***	***PD***	***GBMR***	***SO***	***Proposed***
MSRA	0.29	96.19	3.07	10.36	0.15	23.38	3.37	1.19	5.20
SED	0.22	33.56	1.57	2.00	0.15	5.20	0.85	1.03	2.30
CSSD	0.28	81.58	2.07	2.13	0.14	7.64	0.87	1.11	3.24
Nighttime image	1.62	98.79	25.00	47.85	0.24	151.39	8.72	17.12	38.43

The subjective comparison is shown in Fig. 6 and Fig. 7. From Fig. 6, the saliency maps obtained by the GBMR, SO and the proposed method have a uniform salient region, and the saliency objects are more similar with the ground-truth binary masks. The saliency maps of NP can't clearly distinguish the salient region from their surroundings. The CA and PD method have good detection effects, but the salient objects they detect are not uniform, and their time consumption is very high. The other approaches can not correctly detect the real salient objects under the condition of complicated background. From Fig. 7, it is evident that our model can better detect the salient objects in low contrast images, and is more effective than the others.

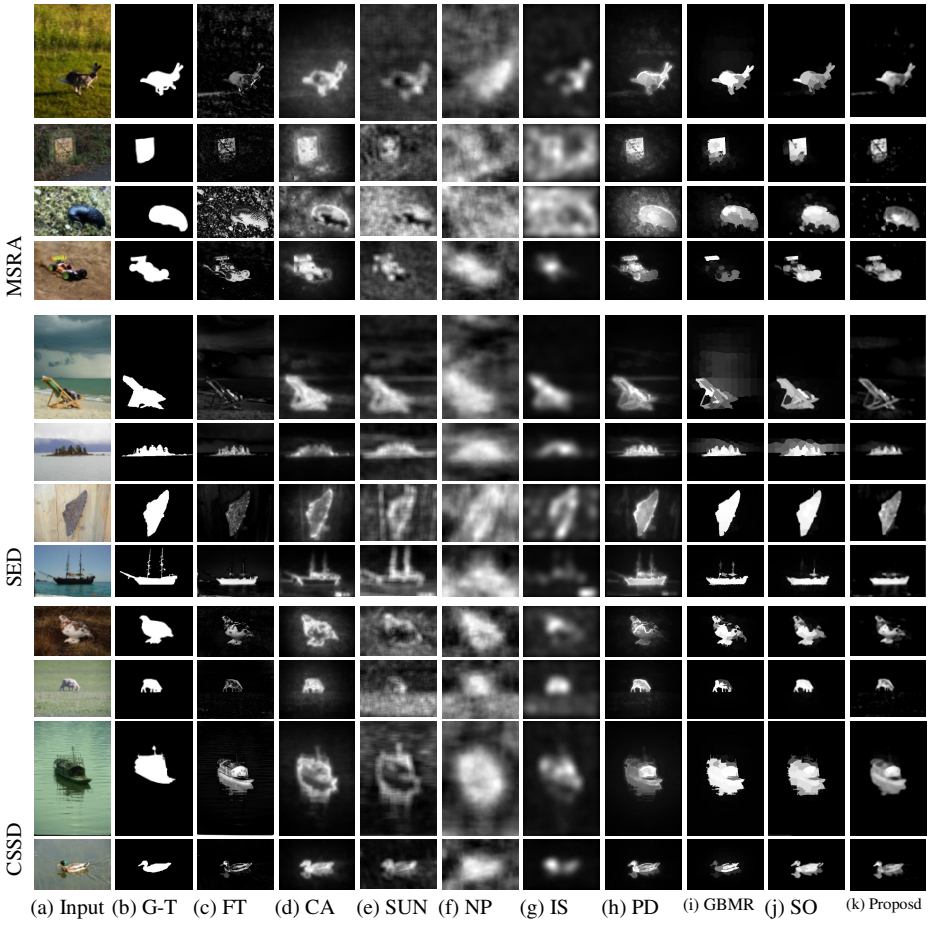


Fig. 6. Qualitative comparisons on MSRA, SED, and CSSD datasets. (a) testing low contrast images, (b) ground-truth binary masks, (c-j) saliency maps obtained by various state-of-the-art saliency models (k) saliency maps obtained by the proposed method.

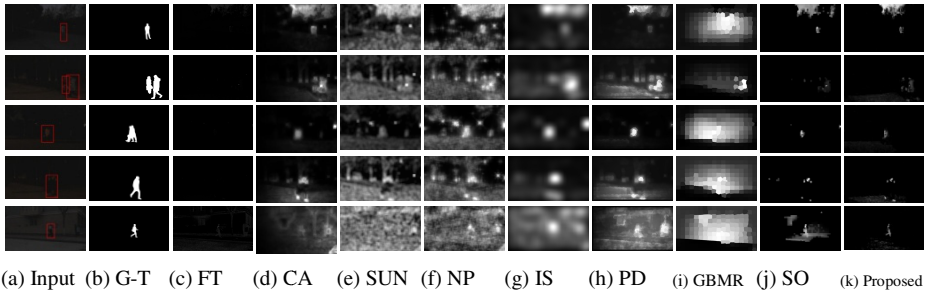


Fig. 7. Qualitative comparison on nighttime image dataset.

4 Conclusions

In this paper, we have proposed an effective superpixel-based saliency model based on the global contrast and the inter-superpixel similarity. Experiments have been carried out on the public available MSRA, SED, CSSD dataset and our nighttime image dataset for salient object detection. Results show that the proposed method outperforms the eight state-of-the-art saliency models. Most of the existing saliency computational methods fail to perform well on low contrast images, while the proposed approach has excellent performance on this task.

Acknowledgments. This work was supported by the Natural Science Foundation of Hubei Provincial of China (2014CFB247) and the National Natural Science Foundation of China (No. 61440016).

References

1. Achanta, R., Hemami, S., Estrada, F., Susstrunk, S.: Frequency-tuned salient region detection. In: IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1597–1604 (2009)
2. Goferman, S., Zelnik-Manor, L., Tal, A.: Context-aware saliency detection. IEEE Trans. on Pattern Analysis and Machine Intelligence **34**(10), 1915–1926 (2012)
3. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. IEEE Trans. Pattern Analysis and Machine Intelligence **20**(11), 1254–1259 (1998)
4. Bruce, N.D.B., Tsotsos, J.K.: Saliency, attention, and visual search: An information theoretic approach. Journal of Vision **9**(3), art. no. 5 (2009)
5. Han, B., Zhu, H., Ding, Y.: Bottom-up saliency based on weighted sparse coding residual. In: Proceedings of the ACM Int. Conf. on Multimedia, pp. 1117–1120 (2011)
6. Zhang, L., Tong, M.H., Marks, T.K., Shan, H., Cottrell, G.W.: SUN: A Bayesian framework for saliency using natural statistics. Journal of Vision **8**(7), art. no. 32 (2008)
7. Rosin, P.L.: A simple method for detecting salient regions. Pattern Recognition **42**(11), 2363–2371 (2009)
8. Cheng, M.-M., Zhang, G.-X., Mitra, N.J., Huang, X., Hu, S.-M.: Global contrast based salient region detection. In: IEEE Conf. on Computer Vision and Pattern Recognition, pp. 409–416 (2011)
9. Borji, A., Itti, L.: Exploiting local and global patch rarities for saliency detection. In: IEEE Conf. on Computer Vision and Pattern Recognition, pp. 478–485 (2011)
10. Wang, L., Huang, K., Huang, Y., Tan, T.: Object detection and tracking for night surveillance based on salient contrast analysis. In: IEEE Int. Conf. on Image Processing, pp. 1113–1116 (2009)
11. Han, J., Yue, J., Zhang, Y., Bai, L.-F.: Salient contour extraction from complex natural scene in night vision image. Infrared Physics & Technology **63**, 165–177 (2014)
12. Liu, T., Sun, J., Zheng, N.-N., Tang, X., Shum, H.-Y.: Learning to detect a salient object. In: IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1–8 (2007)
13. Alpert, S., Galun, M., Basri, R., Brandt, A.: Image segmentation by probabilistic bottom-up aggregation and cue integration. In: IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1–8 (2007)

14. Yan, Q., Xu, L., Shi, J., Jia, J.: Hierarchical saliency detection. In: IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1155–1162 (2013)
15. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. on Pattern Analysis and Machine Intelligence **34**(11), 2274–2282 (2012)
16. Liu, Z., Meur, L., Luo, S.: Superpixel-based saliency detection. In: International Workshop on Image Analysis for Multimedia Interactive Services, pp. 1–4 (2013)
17. Murray, N., Vanrell, M., Otazu, X., Parraga, C.A.: Saliency estimation using a non-parametric low-level vision model. In: IEEE Conf. on Computer Vision and Pattern Recognition, pp. 433–440 (2011)
18. Hou, X., Harel, J., Koch, C.: Image Signature: Highlighting sparse salient regions. IEEE Trans. on Pattern Analysis and Machine Intelligence **34**(1), 194–201 (2012)
19. Margolin, R., Zelnik-Manor, L., Tal, A.: What makes a patch distinct? IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1139–1146 (2013)
20. Yang, C., Zhang, L., Lu, H., Ruan, X., Yang, M.-H.: Saliency detection via graph-based manifold ranking. In: IEEE Conf. on Computer Vision and Pattern Recognition, pp. 3166–3137 (2013)
21. Zhu, W., Liang, S., Wei, Y., Sun, J.: Saliency optimization from robust background detection. In: IEEE Conf. on Computer Vision and Pattern Recognition, pp. 2814–2821 (2014)