

Saliency Detection Using Texture and Local Cues

Qiang Qi^{1,2}, Muwei Jian^{1,2}(✉), Yilong Yin¹, Junyu Dong²,
Wenyin Zhang³, and Hui Yu⁴

¹ School of Computer Science and Technology,
Shandong University of Finance and Economics, Jinan, China
jianmuweihk@163.com

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

³ School of Information Science and Engineering, Linyi University, Linyi, China

⁴ School of Creative Technologies, University of Portsmouth, Portsmouth, UK

Abstract. In this paper, a simple but effective method is proposed for detecting salient objects by utilizing texture and local cues. In contrast to the existing saliency detection models, which mainly consider visual features such as orientation, color, and shape information, our proposed method takes the significant texture cue into consideration to guarantee the accuracy of the detected salient regions. Firstly, an effective method based on selective contrast (SC), which explores the most distinguishable component information in texture, is used to calculate the texture saliency map. Then, we detect local saliency by using a locality-constrained linear coding algorithm. Finally, the output saliency map is computed by integrating texture and local saliency cues simultaneously. Experimental results, based on a widely used and openly available database, demonstrate that the proposed method can produce competitive results and outperforms some existing popular methods.

1 Introduction

Saliency detection is the process of identifying the most informative location of objects in images, which is different to the traditional models of predicting human fixations [44]. In recent years, saliency detection has attracted wide attention of many researchers and become a very active topic in computer vision research. Since it plays a significant role in many computer-vision related applications such as: object detection [11, 15, 17, 36], image/video resizing [3, 35], image/video quality assessment [20, 42], vision tracking [43] etc.

Generally, saliency detection methods can be traditionally summarized into two categories: the bottom-up and top-down approaches. Bottom-up saliency detection methods [37–39] mainly rely on visual-driven stimuli, which directly utilize low-level features for saliency detection. One of the pioneering computational models of bottom-up framework was proposed by Itti et al. in [13]. Later, in order to make salient object detection more precise, a novel graph-based method

was proposed to estimate saliency map in [10]. While the graph-based method usually generates low resolution saliency maps. Then, a spectral residual-based saliency detection method was proposed by Hou and Zhang in [12]. However, the early researches are hard to handle well in the cases with complex scenes. To solve this problem and acquire accurate salient regions, Ma and Zhang [21] presented a local contrast-based method for saliency detection, which increases the accuracy of saliency detection. From then on, many saliency detection models have been proposed and achieved excellent performance in various fields. In [7], a discriminant center-surround hypothesis model was proposed to detect salient objects. In [45], a bottom-up framework by using natural statistics was proposed to detect salient image regions. Achanta et al. [1] proposed a novel saliency detection method based on frequency tuned model to detect salient object. In [23], a new saliency measure was proposed by using a statistical framework and local feature contrast. Wang et al. [29] proposed a novel saliency detection model based on analyzing multiple cues to detect salient objects. In [14], a hierarchical graph saliency detection model was proposed via using concavity context to compute weights between nodes. Tavakoli et al. [25] designed a fast and efficient saliency detection model using sparse sampling and kernel density estimation. In [8], a context-aware based model, which aims at using the image regions to represent the scene, was proposed for saliency detection. Yang et al. [33] presented a new method based on foreground and background cues to achieve a final saliency map. In [32], a novel approach, which can combine contrast, center and smoothness priors, was proposed for salient detection. Ran et al. [24] introduced an effective framework by involving pattern and color distinctness to estimate the saliency value in an image. In [6], a cluster-based method was proposed for co-saliency detection. In [27], a novel saliency measure based on both global and local cues was proposed. Li et al. [19] proposed a visual saliency detection algorithm based on dense and sparse reconstruction. In [22], a novel Cellular Automata (CA) model was introduced to compute the saliency of the objects. In [14], a visual-attention-aware model was proposed for salient-object detection. Recently, a bottom-up saliency-detection method by integrating Quaternionic Distance Based Weber Descriptor (QDWD), center and color cues, was presented in [16]. In [2], a novel and effective deep neural network method incorporating low-level features was proposed for salient object detection.

Top-down based saliency detection models [40, 41], which consider both visual information and prior knowledge, are generally task dependent or application oriented. Thus, the top-down model normally needs supervised learning and is lacking in extendibility. Compared with the bottom-up saliency detection model, not much work has been proposed. Kanan et al. [18] presented an appearance-based saliency model using natural statistics to estimate the saliency of an input image. In [4], a novel top-down framework, which incorporated a tightly coupled image classification module, was proposed for salient object detection. In [5], a weakly supervised top-down approach was proposed for saliency detection by using binary labels.

In this paper, a novel and efficient method based on a bottom-up mechanism, by integrating texture and local cues together, is proposed for saliency detection. In contrast to the existing saliency detection models, we compute the texture saliency map based on selective contrast (SC) method, which can guarantee the accuracy of the detected salient regions. In addition, we incorporated an improved locality-constrained linear coding algorithm (LLC) to detect the local saliency in an image. In order to evaluate the performance of our proposed method, we carried out experiments based on a widely used dataset. Experimental results, compared with other state-of-the-art saliency-detection algorithms, show that our approach is effective and efficient for saliency detection.

The remainder of the paper is organized as follows. In Sect. 2, we introduce the proposed saliency-detection algorithm in detail. In Sect. 3, we demonstrated our experimental results based on a widely used dataset and compared the results with other eight saliency detection methods. The paper closes with a conclusion and discussion in Sect. 4.

2 The Proposed Saliency Detection Method

This section presents the proposed saliency detection method by using selective contrast (SC) method for describing the texture saliency and locality-constrained method for estimating the local saliency. Texture saliency estimation method based on SC will first be described, followed by the locality-constrained method. All these different types of visual cues are fused to form a final saliency map.

2.1 Texture Saliency Detection Based on Selective Contrast

Texture is an important characteristic for human visual perception, which is caused by different physical-reflection properties on the surface of an object with the gray level or color changes. An image is not just a random collection of texture pixels, but a meaningful arrangement of them. Different arrangements of these pixels form different textures would provide us with important saliency information. As a basic property of image, texture also affects the similarity degree among regions, which is a useful cue in saliency detection, so we take the texture cue into consideration in our designed model.

In this section, we adopt selective contrast (SC) based algorithm [30] to estimate the texture saliency. Assume that the pixels in an input image are denoted as $X_n, n = 1, 2, 3, \dots, N$, where N is the number of pixels in the image and X_n is a texture vector, which is achieved by using the uniform LBP [9]. Since the outputs of obtained textures span a very wide range of high dimensional spaces, we hope to express them in a more compact way. In order to solve this issue, we use k -means to cluster these texture expressions and consider the cluster centers as the representative textures. After this transformation, each texture is denoted as its nearest texture prototype, this expression is called selective contrast (SC), which explores the most distinguishable component information

in texture. Thus, each pixels texture saliency based on selective contrast (SC) can be written as follows [30]:

$$S_t(i, R_i) = \sum_{j \subset R_i} d(i, j) \tag{1}$$

where i is the examined pixel, R_i denotes the supporting region for defining the saliency of pixel i , and $d(i, j)$ is the distance of texture descriptors between i and j . The l_2 norm can be used to define the distance measure of textures descriptors:

$$d(i, j) = \|k_i - k_j\| \tag{2}$$

where k_i and k_j are the transformed textures expressions of pixels i and j by $k - means$.

In order to further improve the efficiency by reducing the computational complexity, a limited number of textons are trained from a set of images and the textures can be quantized to M textons. By this means, the computation is reduced to looking up a distance dictionary D_t of $M * M$ dimensions. Thus, the Eq. (1) can be rewritten as:

$$S_t(i, R_i) = \sum_{\psi(j)} f(\psi(j)) D_t(\psi(i), \psi(j)) \tag{3}$$

where i, R_i have a similar meaning in Eq. (1), $\psi(i)$ is the function mapping pixel i to its corresponding prototype texture and $f(\psi(j))$ is the frequency of $\psi(j)$ in region R_i . Readers can be referred to the [30] for more details.

Figure 1(b) shows some texture saliency maps of the input images. From the results we can see that the texture saliency maps can exclude most of the background pixels and detect almost the whole salient objects in images, but it leads to missing some homogeneous regions which have the similar texture appearances.

2.2 Local Saliency Detection Using Locality-Constrained Method

The motivation of local estimation is the local outliers, which are standing out from their neighbors with different colors or textures and tend to attract human attention. In order to detect local outliers and get acceptable performance, local coordinate coding method [34], which described the locality is more essential than sparsity, has been used in saliency detection. Furthermore, the proposed SC model in Sect. 2.1 only takes the texture information into consideration, which misses some local in-formation. Thus, we employ an approximated algorithm based on locality-constrained linear coding (LLC) [28] to estimate the local saliency. For a given image, we first over-segmented the image into N regions, $r_i, i = 1, 2, \dots, N$. For each region r_i , let X be a set of 64-dimensions local descriptors extracted from the image, and $X = [x_i^0, x_i^1, \dots, x_i^{63}]^T, i = 1, 2, \dots, N$. Therefore, the original function of LLC method is written as follows [28]:

$$\min_B \sum_{i=1}^N (\|x_i - D b_i\|^2 + w \|dr_i \cdot * b_i\|^2), s.t. 1^T b_i = 1, \forall_i \tag{4}$$

where $B = [b_1, b_2, \dots, b_N]$ is the set of codes for X , $D = [d_1, d_2, \dots, d_M]$ is the codebook with M entries, and the parameter ω is used to balance the weight between the penalty term and regularization term. The constraint $1^T b_i = 1$ follows the shift-invariant requirements of the LLC coding and $*$ denotes an element-wise multiplication. Here, the dr_i is the locality adaptor that gives different freedom for each codebook vector based on its similarity to the input descriptor x_i , and is defined:

$$dr_i = \exp\left(\frac{\text{dist}(x_i, D)}{\lambda}\right) \quad (5)$$

where $\text{dist}(x_i, D) = [\text{dist}(x_i, d_1), \text{dist}(x_i, d_2), \dots, \text{dist}(x_i, d_M)]^T$, $\text{dist}(x_i, d_i)$ denotes the Euclidean distance between x_i and the code book vector d_i , λ is used to adjust the weight decay speed for the locality adaptor, and M is the number of elements in the codebook. More details about LLC can be referred to [28, 34].

In this paper, we adopted an approximated LLC algorithm [27] to detect the local cue. We consider the K nearest neighbors in spatial as the local basis D_i owing to the vector b_i in Eq. (4) with a few non-zero values, which means that it is sparse in some extent. It should be noted that the K is smaller than the size of the original codebook M . Thus, the Eq. (4) can be rewritten as follows:

$$\min_B \left(\sum_{i=1}^N \|x_i - D b_i\|^2 \right), \text{ s.t. } 1^T b_i = 1, \forall_i \quad (6)$$

where D_i denotes the new codebook for each region r_i , $i = 1, 2, \dots, N$ and K is the size of the new codebook and is empirically set at $K = 2M/3$.

Unlike the traditional LLC algorithm, solving the improved LLC algorithm is simple and the solution can be derived analytically by

$$b_i = \frac{1}{(C_i + \omega * \text{dig}(C_i))}, \quad (7)$$

$$\tilde{b}_i = \frac{b_i}{1^T b_i}, \quad (8)$$

where $C_i = (D_i - 1x_i^T)(D_i - 1x_i^T)^T$ represents the covariance matrix of the feature and ω is a regularization parameter, which is set to be 0.1 in the proposed algorithm. As the solution of the improved LLC method is simple and fast, therefore, the local saliency value of the region r_i can be defined as follows [27]:

$$S(r_i) = \|x_i - D_i \tilde{b}_i\|^2, \quad (9)$$

where \tilde{b}_i is the solution of Eq. (6), which is achieved by Eqs. (7) and (8).

Figure 1(c) shows some local saliency maps. We can see that the local saliency maps achieve more reliable local detailed information owing to the locality-constrained coding model.

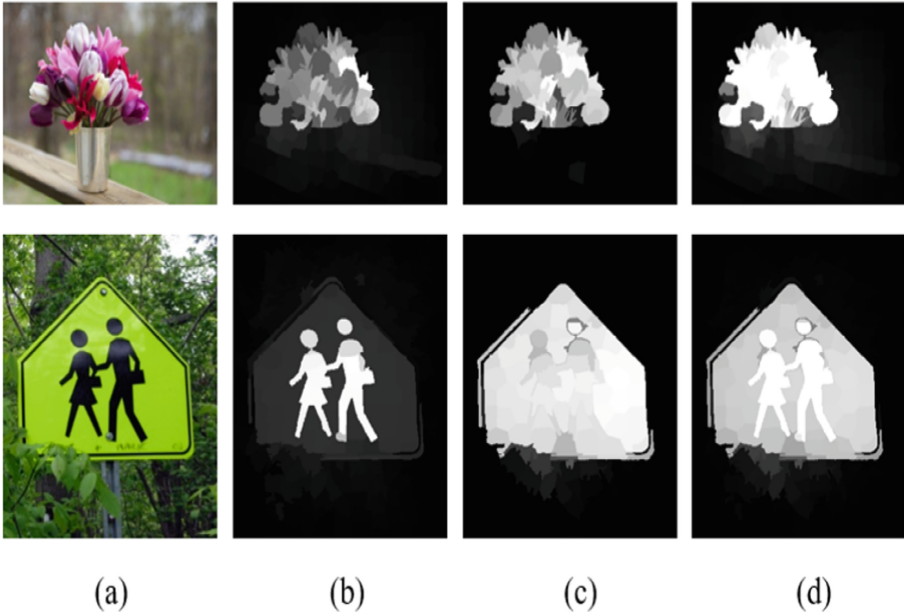


Fig. 1. The proposed method for saliency detection. (a) Input images, (b) texture saliency maps, (c) local saliency maps, (d) the final saliency maps.

2.3 Final Saliency Fusion

The texture saliency map and the local saliency map are linearly combined with adaptive weights to define the final saliency map. Therefore, the final saliency map can be defined as follows:

$$Sal = \alpha S_t + \beta S_l, \quad (10)$$

where S_t and S_l are the texture saliency and the local saliency, respectively. α and β are the weights of the texture saliency and the local saliency, accordingly. In this paper, we introduced a more effective and logical fusion method to adjust the weights between the different feature maps adaptively. We used the *DOS* (degree-of-scattering) of saliency map to determine the weighting parameters and set the $\alpha = 1 - DOS^t$, $\beta = 1 - DOS^t$ according to the method in [26].

3 Experimental Results

We perform saliency detection experiments based on a widely used dataset: ECSSD [31], which included 1000 images acquired from the internet, and compare our method with other eight state-of-the-art methods including the Spectral Residual (SR) [12], Saliency detection using Natural Statistics (SUN) [45], Segmenting salient objects (SEG) [23], Context-aware (CA) saliency detection [8],

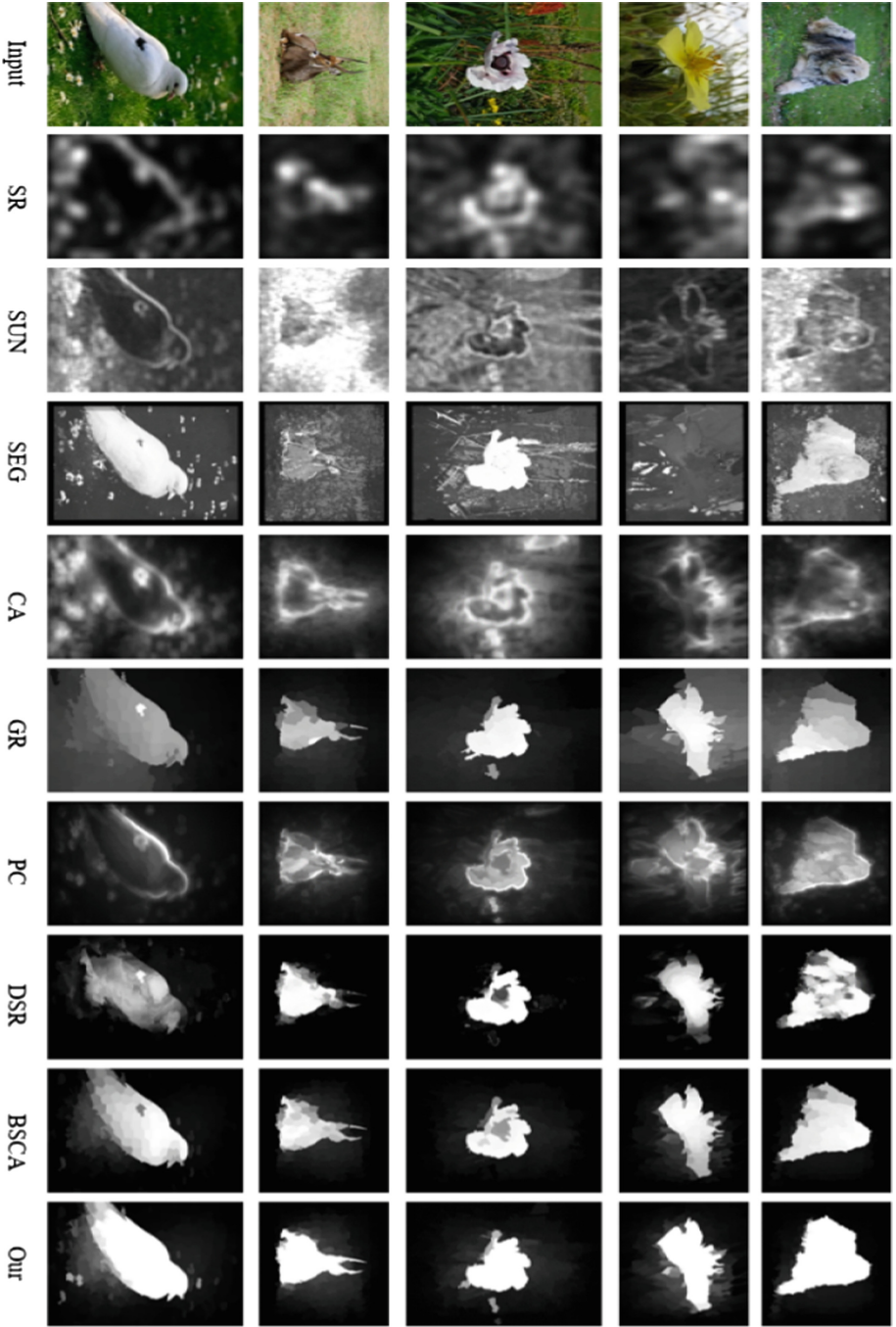


Fig. 2. Qualitative comparisons of different approaches based on ECSSD database.

Graph-regularized (GR) saliency detection [32], Pattern distinctness and color (PC) based method [24], Dense and sparse reconstruction (DSR) [19], Background-Single Cellular Automata (BSCA) based saliency detection [22].

In order to quantitatively compare the state-of-the-art saliency-detection methods, the average precision, recall, and are utilized to measure the quality of the saliency maps based on setting a segmentation threshold for binary segmentation. The adaptive threshold is twice the average value of the whole saliency map to get the accurate results. Each image is segmented with superpixels and masked out when the mean saliency values are lower than the adaptive threshold. The is defined as follows:

$$F_{\beta} = \frac{(1 + \beta) * Precision * Recall}{\beta * Precision + Recall}, \tag{11}$$

where β is a real positive value and is set at $\beta = 0.3$.

We compared the performance of the proposed method with other eight state-of-the-art saliency detection methods. Figure 2 shows some saliency detection results of different methods based on the ECSSD database. The output results

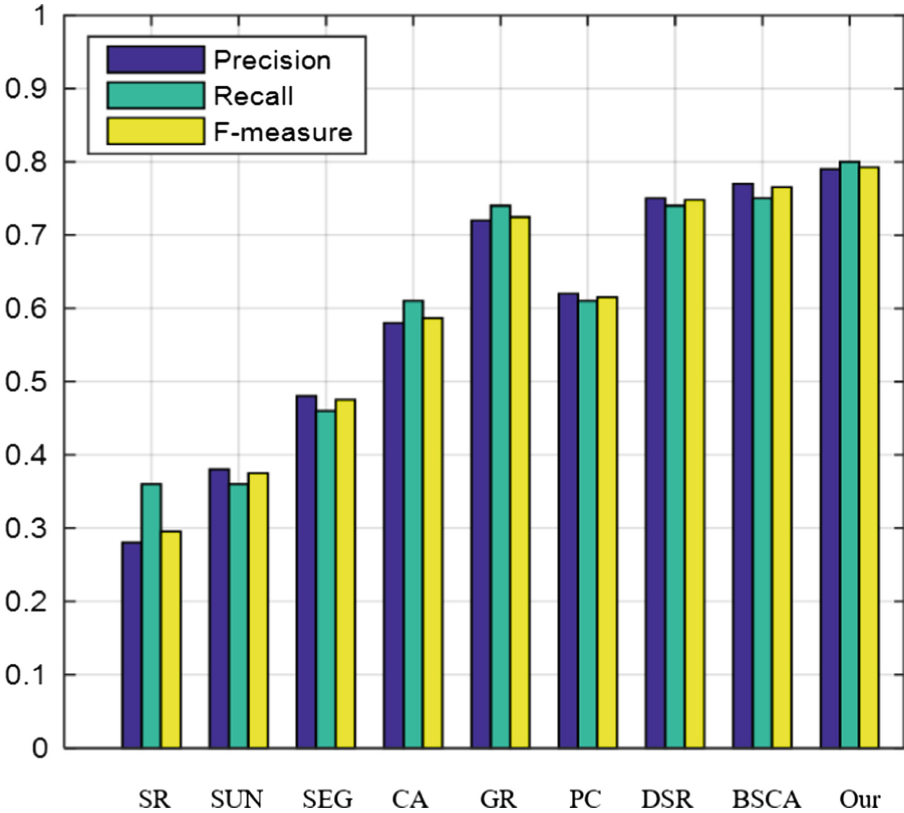


Fig. 3. Comparison of different saliency-detection methods in terms of average precision, recall, and based on ECSSD database.

show that our final saliency maps can accurately detect almost entire salient objects and preserve the salient objects contours more clearly.

We also used the precision, recall and the $F - measure$ to evaluate the performance of different methods objectively. Figure 3 shows the comparisons of different methods under different evaluation criterions. As can be seen from Fig. 3, our pro-posed method outperforms the other eight methods in terms of detection accuracy and the proposed method achieves the best overall saliency-detection performance (with precision = 79.0%, recall = 80.0%), and the $F - measure$ is 79.2%. The experiment results show that the proposed model is efficient and effective.

4 Conclusion and Discussion

This paper proposes a novel bottom-up method for efficient and accurate image saliency detection. This proposed approach integrated texture and local cues to estimate the final saliency map. The texture saliency maps are computed based on the selective contrast (SC) method, which explores the most distinguishable component information in texture. The local saliency maps are achieved by utilizing a locality-constrained method. We also evaluated our method based on a publicly available dataset and compared our proposed approach with other eight different state-of-the-art methods. The experimental results show that our algorithm can produce promising results compared to the other state-of-the-art saliency-detection models.

Acknowledgments. This work was supported by National Natural Science Foundation of China (NSFC) (61601427); Natural Science Foundation of Shandong Province (ZR2015FQ011); Applied Basic Research Project of Qingdao (16-5-1-4-jch); China Postdoctoral Science Foundation funded project (2016M590659); Postdoctoral Science Foundation of Shandong Province (201603045); Qingdao Postdoctoral Science Foundation funded project (861605040008) and The Fundamental Research Funds for the Central Universities (201511008, 30020084851).

References

1. Achanta, R., Hemami, S., Estrada, F., Susstrunk, S.: Frequency-tuned salient region detection. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009, pp. 1597–1604 (2009)
2. Chen, J., Chen, J., Lu, H., Chi, Z.: CNN for saliency detection with low-level feature integration. *Neurocomputing* **226**(C), 212–220 (2017)
3. Cheng, M.-M., Liu, Y., Hou, Q., Bian, J., Torr, P., Hu, S.-M., Tu, Z.: HFS: hierarchical feature selection for efficient image segmentation. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9907, pp. 867–882. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46487-9_53
4. Cholakkal, H., Johnson, J., Rajan, D.: A classifier-guided approach for top-down salient object detection. *Sig. Process. Image Commun.* **45**(C), 24–40 (2016)
5. Cholakkal, H., Johnson, J., Rajan, D.: Weakly supervised top-down salient object detection (2016)

6. Fu, H., Cao, X., Tu, Z.: Cluster-based co-saliency detection. *IEEE Trans. Image Process.* **22**(10), 3766 (2013). A Publication of the IEEE Signal Processing Society
7. Gao, D., Mahadevan, V., Vasconcelos, N.: The discriminant center-surround hypothesis for bottom-up saliency. In: *Advances in Neural Information Processing Systems*, vol. 20, pp. 497–504 (2007)
8. Goferman, S., Zelnikmanor, L., Tal, A.: Context-aware saliency detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(10), 1915–1926 (2012)
9. Guo, Z., Zhang, L., Zhang, D.: Rotation invariant texture classification using LBP variance (LBPV) with global matching. *Pattern Recogn.* **43**(3), 706–719 (2010)
10. Harel, J., Koch, C., Perona, P.: Graph-based visual saliency. In: *International Conference on Neural Information Processing Systems*, pp. 545–552 (2006)
11. Hou, Q., Cheng, M.M., Hu, X.W., Borji, A., Tu, Z., Torr, P.: Deeply supervised salient object detection with short connections (2017)
12. Hou, X., Zhang, L.: Saliency detection: a spectral residual approach. In: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2007*, pp. 1–8 (2007)
13. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(11), 1254–1259 (1998). IEEE Computer Society
14. Jian, M., Lam, K.M., Dong, J., Shen, L.: Visual-patch-attention-aware saliency detection. *IEEE Trans. Cybern.* **45**(8), 1575–1586 (2015)
15. Jian, M., Qi, Q., Dong, J., Sun, X., Sun, Y., Lam, K.-M.: The OUC-vision large-scale underwater image database (2017)
16. Jian, M., Qi, Q., Dong, J., Sun, X., Sun, Y., Lam, K.M.: Saliency detection using quaternionic distance based weber descriptor and object cues. In: *Signal and Information Processing Association Summit and Conference*, pp. 1–4 (2017)
17. Jian, M., Qi, Q., Dong, J., Sun, X., Sun, Y., Lam, K.-M.: Saliency detection using quaternionic distance based weber local descriptor and level priors. *Multimed. Tools Appl.* (2017). <https://doi.org/10.1007/s11042-017-5032-z>
18. Kanan, C., Tong, M.H., Zhang, L., Cottrell, G.W.: SUN: top-down saliency using natural statistics. *Vis. Cogn.* **17**(6–7), 979 (2009)
19. Li, X., Lu, H., Zhang, L., Xiang, R., Yang, M.H.: Saliency detection via dense and sparse reconstruction. In: *IEEE International Conference on Computer Vision*, pp. 2976–2983 (2013)
20. Ma, Q.: New strategy for image and video quality assessment. *J. Electron. Imaging* **19**(1), 011019 (2010)
21. Ma, Y.F., Zhang, H.J.: Contrast-based image attention analysis by using fuzzy growing. In: *Eleventh ACM International Conference on Multimedia*, pp. 374–381 (2003)
22. Qin, Y., Lu, H., Xu, Y., Wang, H.: Saliency detection via cellular automata. In: *Computer Vision and Pattern Recognition*, pp. 110–119 (2015)
23. Rahtu, E., Kannala, J., Salo, M., Heikkilä, J.: Segmenting salient objects from images and videos. In: *Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010. LNCS, vol. 6315*, pp. 366–379. Springer, Heidelberg (2010). <https://doi.org/10.1007/978-3-642-15555-0-27>
24. Ran, M., Tal, A., Zelnikmanor, L.: What makes a patch distinct? In: *Computer Vision and Pattern Recognition*, pp. 1139–1146 (2013)
25. Rezazadegan Tavakoli, H., Rahtu, E., Heikkilä, J.: Fast and efficient saliency detection using sparse sampling and kernel density estimation. In: *Heyden, A., Kahl, F. (eds.) SCIA 2011. LNCS, vol. 6688*, pp. 666–675. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21227-7_62

26. Tian, H., Fang, Y., Zhao, Y., Lin, W.: Salient region detection by fusing bottom-up and top-down features extracted from a single image. *IEEE Trans. Image Process.* **23**(10), 4389–4398 (2014). A Publication of the IEEE Signal Processing Society
27. Tong, N., Lu, H., Zhang, Y., Xiang, R.: Salient object detection via global and local cues. *Pattern Recogn.* **48**(10), 3258–3267 (2015)
28. Wang, J., Yang, J., Yu, K., Lv, F., Huang, T., Gong, Y.: Locality-constrained linear coding for image classification, vol. 119, no. 5, pp. 3360–3367 (2010)
29. Wang, L., Xue, J., Zheng, N., Hua, G.: Automatic salient object extraction with contextual cue, vol. 23, no. 5, pp. 105–112 (2011)
30. Wang, Q., Yuan, Y., Yan, P.: Visual saliency by selective contrast. *IEEE Trans. Circ. Syst. Video Technol.* **23**(7), 1150–1155 (2013)
31. Yan, Q., Xu, L., Shi, J., Jia, J.: Hierarchical saliency detection. In: *Computer Vision and Pattern Recognition*, pp. 1155–1162 (2013)
32. Yang, C., Zhang, L., Lu, H.: Graph-regularized saliency detection with convex-hull-based center prior. *IEEE Signal Process. Lett.* **20**(7), 637–640 (2013)
33. Yang, C., Zhang, L., Lu, H., Ruan, X., Yang, M.H.: Saliency detection via graph-based manifold ranking. In: *Computer Vision and Pattern Recognition*, pp. 3166–3173 (2013)
34. Yu, K., Zhang, T., Gong, Y.: Nonlinear learning using local coordinate coding. In: *International Conference on Neural Information Processing Systems*, pp. 2223–2231 (2009)
35. Zhang, J.: Seam carving for content-aware image resizing. *ACM Trans. Graph.* **26**(3), 10 (2007)
36. Song, H., Liu, Z., Du, H., et al.: Depth-aware salient object detection and segmentation via multiscale discriminative saliency fusion and bootstrap learning. *IEEE Trans. Image Process.* **26**(9), 4204–4216 (2017)
37. Guan, Y., Jiang, B., Xiao, Y., et al.: A new graph ranking model for image saliency detection problem. In: *Software Engineering Research, Management and Applications (SERA)*, pp. 151–156 (2017)
38. He, Z., Jiang, B., Xiao, Y., Ding, C., Luo, B.: Saliency detection via a graph based diffusion model. In: Foggia, P., Liu, C.-L., Vento, M. (eds.) *GbrPR 2017*. LNCS, vol. 10310, pp. 3–12. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58961-9_1
39. Peng, H., Li, B., Ling, H., et al.: Salient object detection via structured matrix decomposition. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(4), 818–832 (2017)
40. Yang, J., Yang, M.H.: Top-down visual saliency via joint CRF and dictionary learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(3), 576–588 (2017)
41. Deng, T., Yang, K., Li, Y., et al.: Where does the driver look? top-down-based saliency detection in a traffic driving environment. *IEEE Trans. Intell. Transp. Syst.* **17**(7), 2051–2062 (2016)
42. Liu, Y., Yang, J., Meng, Q., et al.: Stereoscopic image quality assessment method based on binocular combination saliency model. *Signal Process.* **125**, 237–248 (2016)
43. Zhang, K., Liu, Q., Wu, Y.: Robust visual tracking via convolutional networks without training. *IEEE Trans. Image Process.* **25**(4), 1779–1792 (2016)
44. Lee, S.H., Kang, J.W., Kim, C.S.: Compressed domain video using global and local spatiotemporal features. *J. Vis. Commun. Image Represent.* **35**, 169–183 (2016)
45. Zhang, L., Tong, M.H., Marks, T.K., Shan, H., Cottrell, G.W.: Sun: a Bayesian framework for saliency using natural statistics. *J. Vis.* **8**(7), 32 (2008)