Saliency Detection for Stereoscopic Images Based on Depth Confidence Analysis and Multiple Cues Fusion

Runmin Cong, *Student Member, IEEE*, Jianjun Lei, *Member, IEEE*, Changqing Zhang, Qingming Huang, *Senior Member, IEEE*, Xiaochun Cao, *Senior Member, IEEE*, and Chunping Hou

***Abstract*—Stereoscopic perception is an important part of human visual system that allows the brain to perceive depth. However, depth information has not been well explored in existing saliency detection models. In this letter, a novel saliency detection method for stereoscopic images is proposed. First, we propose a measure to evaluate the reliability of depth map, and use it to reduce the influence of poor depth map on saliency detection. Then, the input image is represented as a graph, and the depth information is introduced into graph construction. After that, a new definition of compactness using color and depth cues is put forward to compute the compactness saliency map. In order to compensate the detection errors of compactness saliency when the salient regions have similar appearances with background, foreground saliency map is calculated based on depth-refined foreground seeds' selection (DRSS) mechanism and multiple cues contrast. Finally, these two saliency maps are integrated into a final saliency map through weighted-sum method according to their importance. Experiments on two publicly available stereo data sets demonstrate that the proposed method performs better than other ten state-of-the-art approaches.**

***Index Terms*—Color and depth-based compactness, depth confidence measure, multiple cues, saliency detection.**

I. Introduction

ACCORDING to the principle of human visual perception, people would like to place more attention on the region that stands out from the image background. Saliency detection is becoming an increasingly important branch in computer vision due to its various applications in object detection and recognition [1], [2], image retrieval [3], image compression [4], [5], and image retargeting [6], [7]. In general, saliency detection models can be categorized into data-driven bottom-up model and task-driven top-down model [8]. In this letter, we focus on bottom-up saliency detection models.

Manuscript received January 30, 2016; revised March 14, 2016; accepted April 12, 2016. Date of publication April 15, 2016; date of current version May 05, 2016. This work was supported by the Natural Science Foundation of China under Grant 61271324, Grant 61520106002, Grant 61471262, and Grant 91320201. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Alexandre X. Falcao. *(Corresponding author: Jianjun Lei.)*

R. Cong, J. Lei, and C. Hou are with the School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China (e-mail: rmcong@tju.edu.cn; jjlei@tju.edu.cn; hcp@tju.edu.cn).

C. Zhang is with the School of Computer Science and Technology, Tianjin University, Tianjin 300072, China (e-mail: zhangchangqing@tju.edu.cn).

Q. Huang is with the Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China (e-mail: qmhuang@ucas.ac.cn).

X. Cao is with the Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China (e-mail: caoxiaochun@iie.ac.cn).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2016.2557347

Saliency detection aims to effectively highlight salient regions and suppress background regions. By far, many saliency detection methods for RGB image have been put forward, which integrate different visual cues to compute the saliency map. As a pioneer, Itti *et al.* [9] presented a multiscale saliency detection model that computes center–surround differences using three visual features including color, intensity, and orientation. Cheng *et al.* [10], [11] proposed a region-based saliency model that measures the global contrast between the target regions with respect to all other regions in the image. In [12], graph-based manifold ranking was introduced into saliency detection model, which ranks the differences between salient object and background. Wei *et al.* [13] proposed a multiple saliency methods' fusion framework based on Dempster–Shafer theory (DST), and improved the performance of saliency detection. Sun *et al.* [14] cast the saliency detection into a Markov problem, which integrate background prior and Markov absorption probability on a weighted graph. Recently, some new methods combine multiple cues into saliency detection model to achieve a better result. In [15], Zhou *et al.* found that compactness and local contrast are complementary to each other, and proposed a saliency detection method that integrates compactness and local contrast. In [16], Li *et al.* proposed a novel label propagation method for saliency detection, which integrates backgroundness, objectness, and compactness cues.

Most previous works on saliency detection are focused on 2-D images, which mainly concentrate on RGB color information while ignoring depth/disparity cue [17], [18]. In fact, 3-D visual information can supply a useful cue for saliency detection [19], [20]. Niu *et al.* [21] proposed a saliency model for stereoscopic images based on global disparity contrast and domain knowledge in stereoscopic photography. Desingh *et al.* [22] computed the stereo saliency by fusing depth saliency with 2-D saliency models, which through nonlinear support vector regression. In [23], the disparity maps are used to refine the 2-D saliency model and maintain the consistency between the stereo matching and saliency maps. Ju *et al.* [24], [25] proposed a depth-aware method for saliency detection using an anisotropic center–surround difference (ACSD) measure. In addition, they construct a large data set for stereo saliency detection, which includes 1985 stereo images and estimated depth maps.

Considering the effectivity of integrating depth information, we propose a novel saliency detection model for stereoscopic images in this letter. The main contributions can be summarized as follows: 1) a good depth map can be benefit for the saliency detection no matter how it produced. Consequently, it is vital that construct a measure to describe the quality of

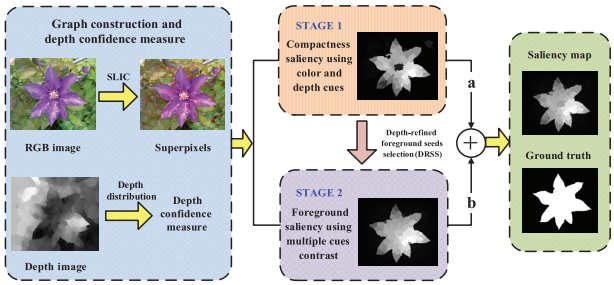

Fig. 1. Flowchart of the proposed method.

Fig. 2. Different qualities of depth maps. (a) Good depth map $\lambda_d = 0.8014$. (b) Common depth map $\lambda_d = 0.3890$. (c) Poor depth map $\lambda_d = 0.0422$.

depth map. According to the observation of depth distribution, a confidence measure for depth map is proposed to reduce the influence of poor depth map on saliency detection; 2) a novel model for compactness integrating color and depth information is put forward to compute the compactness saliency; and 3) a foreground seeds' selection mechanism based on depth refined is presented. The saliency is measured by contrast between the target regions with seed regions, which integrate color, depth, and texture cues.

II. Proposed Method

The flowchart of the proposed stereo saliency detection method is depicted as in Fig. 1. There are four parts in our method. First, a novel confidence measure is calculated to evaluate the reliability of depth map, and used in the following processes. The depth confidence measure can reduce the influence of poor depth map on saliency detection. Simultaneously, RGB image is abstracted into superpixels and represented as a graph. Then, compactness saliency based on color and depth cues is calculated. Furthermore, foreground seeds are selected by combining compactness saliency result and depth map. Taking color, depth, and texture cues into consideration, a multiple cues contrast saliency detection method based on foreground seeds is proposed. Finally, compactness and foreground saliency map are weighted to obtain the final saliency map.

A. Depth Confidence Measure

The quality of depth map is very important for the using of depth cue. A good depth map can provide accurate depth information that benefits for saliency detection. In contrast, the poor depth map may cause a wrong detection. Thus, a depth confidence measure is proposed in this letter to evaluate the reliability of depth map. We found that a good depth map often owns clear hierarchy, and the salient object can be highlighted from the background. Fig. 2 shows some examples of depth maps with different qualities.

We rank the input depth map into three grades roughly, namely, good, common, and bad. Based on the observation of depth statistical characteristic, we found that the values are usually concentrate on lower part for good depth map, whereas the distribution of poor depth map tends to concentrate on relatively larger values. Therefore, the mean value of the whole depth image is an effective parameter to tell them apart. In statistics, coefficient of variation is used to evaluate the dispersion degree of the data. It is observed that the poor depth map appears strong concentration compared with other cases. Thus, the coefficient of variation is introduced in our confidence measure. In addition, there is a more random distribution for a common depth map; therefore, the depth frequency entropy is defined to evaluate the randomness of an input depth image. According to these observations of depth distribution, we define the depth confidence measure as follows:

$$\lambda_d = \exp((1 - m_d) \cdot \mathrm{CV} \cdot H) - 1 \tag{1}$$

where m_d is the mean value of the whole depth image, $\mathrm{CV} = m_d/\sigma_d$ is the coefficient of variation, σ_d is the standard deviation for depth image, and H is the depth frequency entropy, which denotes the randomness of depth distribution, and can be defined as follows:

$$H = -\sum_{i=1}^{L} P_i \log(P_i) \tag{2}$$

where $P_i = n_i/n_\Sigma$, n_Σ is the number of pixels in the depth map, n_i is the number of pixels that belong to the region level r_i, and L is the levels of depth map. In this letter, the input depth map is normalized to $[0, 1]$ firstly. Then, $L - 1$ thresholds, namely, T_k are used to divide the depth map into L levels. A larger λ_d corresponds to more reliable of the input depth map.

B. Graph Construction

The input RGB image is abstracted into homogenous and compact regions using SLIC superpixel segmentation method [26]. The number of superpixel N is set to 200 in experiments. Then, we construct a graph $G = (V, E)$, where V represents the set of nodes, which corresponds to the superpixels generated by SLIC method, and E denotes the set of links between adjacent nodes.

In this work, the Euclidean distance l_{ij} in *CIE Lab* color space and depth difference d_{ij} between superpixels v_i and v_j are defined as

$$l_{ij} = \|\boldsymbol{c}_i - \boldsymbol{c}_j\| \tag{3}$$

and

$$d_{ij} = |d_i - d_j| \tag{4}$$

where $\boldsymbol{c}_i$ is the mean color value of superpixel v_i and d_i denotes the mean depth value of superpixel v_i. The similarity between two superpixels a_{ij} is defined as

$$a_{ij} = \exp\left(-\frac{l_{ij} + \lambda_d \cdot d_{ij}}{\sigma^2}\right) \tag{5}$$

where σ^2 is a parameter to control strength of the similarity, and it is set to 0.1 in all our experiments. The affinity

matrix $\boldsymbol{W} = [w_{ij}]_{N\times N}$ is defined as the similarity between two adjacent superpixels

$$w_{ij} = \begin{cases} a_{ij}, & \text{if } j \in \Omega_i \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

where Ω_i is the set of neighbors of superpixel v_i.

C. Compactness Saliency Using Color and Depth Cues

Intuitively, salient regions have compact spread in spatial domain, whereas the colors of background regions have a larger spread over the whole image [15]. In addition, the values of depth exhibit limited compactness, i.e., the depth values of salient regions are more likely to have a centralized distribution near the center of image. Motivated by this, we calculate the compactness saliency using color and depth cues. To obtain more precise result, we first propagate the similarity a_{ij} using manifold ranking method [27].

Zhou *et al.* proposed a compactness metric using color spatial information in [15]. In this letter, a novel measure is proposed to calculate the color- and depth-based compactness. The saliency map based on compactness is defined as

$$S_{\mathrm{CS}}(i) = [1 - \mathrm{norm}(\mathrm{cc}(i) + \mathrm{dc}(i))] \cdot \mathrm{Obj}(i) \tag{7}$$

where $\mathrm{cc}(i)$ is the color-based compactness of superpixel v_i, $\mathrm{dc}(i)$ is the depth-based compactness of superpixel v_i, and $\mathrm{norm}(x)$ is a function that normalizes x to $[0, 1]$ using min–max normalization. Considering the importance of location for saliency detection, objectness measure [28] $\mathrm{Obj}(i)$ is introduced in our model to evaluate the probability of superpixel v_i that belongs to an object. The color- and depth-based compactness are defined as

$$\mathrm{cc}(i) = \frac{\sum_{j=1}^{N} a_{ij} \cdot n_j \cdot \|\boldsymbol{b}_j - \boldsymbol{\mu}_i\|}{\sum_{j=1}^{N} a_{ij} \cdot n_j} \tag{8}$$

and

$$\mathrm{dc}(i) = \frac{\sum_{j=1}^{N} a_{ij} \cdot n_j \cdot \|\boldsymbol{b}_j - \boldsymbol{p}\| \cdot \exp(-\frac{\lambda_d \cdot d_i}{\sigma^2})}{\sum_{j=1}^{N} a_{ij} \cdot n_j} \tag{9}$$

where n_j denotes the number of pixels that belong to superpixel v_j, which emphasizes the impact of larger region, $\boldsymbol{b}_j = [b_j^x, b_j^y]$ is the centroid coordinate of superpixel v_j, $\boldsymbol{p} = [c_x, c_y]$ is the spatial position of the image center, and the spatial mean $\boldsymbol{\mu}_i = [\mu_i^x, \mu_i^y]$ is defined as

$$\mu_i^x = \frac{\sum_{j=1}^{N} a_{ij} \cdot n_j \cdot b_j^x}{\sum_{j=1}^{N} a_{ij} \cdot n_j} \tag{10}$$

and

$$\mu_i^y = \frac{\sum_{j=1}^{N} a_{ij} \cdot n_j \cdot b_j^y}{\sum_{j=1}^{N} a_{ij} \cdot n_j}. \tag{11}$$

D. Foreground Saliency Using Multiple Cues Contrast

Although compactness saliency detection method is active on some level, there are some limitations. When the salient regions have similar appearances with background, the regions may be wrongly detected. Hence, a foreground saliency detection method based on multiple cues contrast is proposed to mitigate this problem.

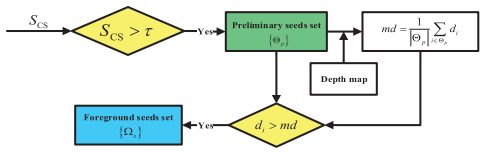

Fig. 3. Flowchart of DRSS mechanism.

Traditionally, foreground seeds are selected just based on the preliminary saliency map. Considering the importance of depth information, a novel selection mechanism for foreground seeds is introduced, in which the seeds should own larger values of compactness saliency and depth simultaneously. Therefore, a depth-refined foreground seeds' selection (DRSS) method is proposed, and the flowchart is shown in Fig. 3. First, preliminary seeds are determined by thresholding segmentation using compactness saliency map, where the threshold τ is set to 0.5. Then, the mean depth value of preliminary seeds is used to refine the preliminary seeds, and obtain the final foreground seeds' set.

Next, we calculate the contrast of each superpixel with the foreground seeds based on multiple cues, which include color, depth, texture, and position information. A superpixel is more likely to be salient if it is more similar to the foreground seeds. The foreground saliency is computed as follows:

$$S_{\mathrm{fg}}(i) = \sum_{j\in\Omega_s} [a_{ij} \cdot D_t(i,j) \cdot \exp(-\|\boldsymbol{b}_i - \boldsymbol{b}_j\|/\sigma^2) \cdot n_j] \tag{12}$$

where Ω_s is the set of foreground seeds, $\|\boldsymbol{b}_i - \boldsymbol{b}_j\|$ denotes the Euclidean distance between position of superpixels, and $D_t(i,j)$ is the texture similarity between superpixels using LBP feature [29], which is defined as follows:

$$D_t(i,j) = \frac{|\boldsymbol{k}_i^T \boldsymbol{k}_j|}{\|\boldsymbol{k}_i\|\|\boldsymbol{k}_j\|} \tag{13}$$

where $\boldsymbol{k}_i$ is the LBP histogram frequency of superpixel v_i. To avoid the problem that saliency map highlights object boundaries rather than the entire region, manifold ranking method is used to propagate the foreground saliency map. Finally, the map after propagation is normalized to $[0, 1]$, and the final foreground saliency map S_{FS} is obtained.

E. Saliency Map Integration

The compactness and foreground saliency maps are complementary to each other. Considering the foreground saliency map is based on the compactness saliency result, we integrate these two saliency map into a final saliency map through weighted-sum method

$$S = \gamma \cdot S_{\mathrm{CS}} + (1-\gamma) \cdot S_{\mathrm{FS}} \tag{14}$$

where γ balances the compactness saliency map S_{CS} and foreground saliency map S_{FS}.

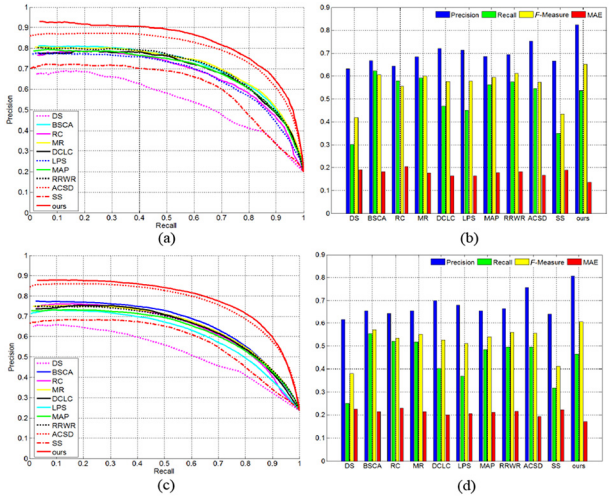

Fig. 4. Quantitative comparison of the proposed method with ten state-of-the-art methods. (a) Precision–recall curves of different methods on NJU-400 data set. (b) Average precision, recall, F-measure, and MAE of different methods on NJU-400 data set. (c) Precision–recall curves of different methods on NJU-1985 data set. (d) Average precision, recall, F-measure, and MAE of different methods on NJU-1985 data set.

III. Experimental Results

We evaluate the performance of our method on two data sets: NJU-400 [24] and NJU-1985 [25], which include RGB images, depth maps, and pixel-wise ground truth annotations. The performance is evaluated using precision-recall curve, F-measure, and mean absolute error (MAE) [10], [30], [31]. The precision–recall curve is obtained by binarizing the saliency map using thresholds in the range of 0 and 255. In all experiments, we set the parameters $L = 3$, $T_1 = 0.4$, $T_2 = 0.6$, and $\gamma = 0.8$.

A. Performance Comparison

We compare our method with eight state-of-the-art 2-D methods: RC [11], MR [12], DS [13], MAP [14], DCLC [15], LPS [16], BSCA [31], and RRWR [32], and two stereo saliency detection methods: SS [21] and ACSD [25]. Fig. 4 shows the evaluation results of the proposed method with ten state-of-the-art methods on two data sets. On both data sets, the precision–recall curves show that the proposed method performed better than other methods. Similarly, the proposed method achieves the best performance in terms of the average precision, F-measure, and MAE compared with other approaches due to the depth confidence analysis and two-stage saliency computation mechanism. Taking the F-measure on the NJU-1985 data set as an example, the F-measure values of two stereo saliency detection methods (ACSD and our method) are 0.5552 and 0.6055, respectively. Nevertheless, our method achieved a lightly lower average recall value than ACSD method. In the future work, the recall value should be improved on the premise of that the value of precision has been maintained. Fig. 5 presents the visual comparison of different saliency detection methods. The proposed method has more similar appearances with ground truth, and owns clear contour and uniform salient regions. Furthermore, the F-measure values in Fig. 5 also demonstrate the effectiveness of our proposed method.

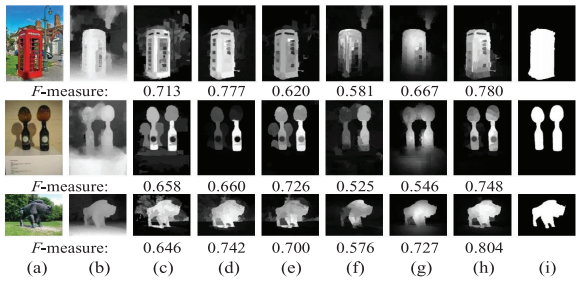

Fig. 5. Visual comparison of saliency maps. (a) Input RGB image. (b) Input depth image. (c) RC [11]. (d) RRWR [32]. (e) DCLC [15]. (f) SS [21]. (g) ACSD [25]. (h) Ours. (i) Ground truth.

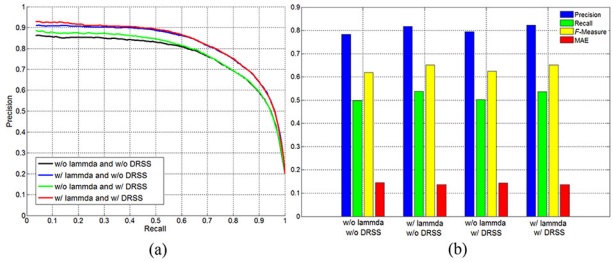

Fig. 6. Evaluation of different factors for final saliency detection on NJU-400 data set. (a) Precision–recall curves of different factors. (b) Average precision, recall, F-measure, and MAE of different factors.

B. Parameter Analysis

In this section, we evaluate the performance of our method under different factors. The precision–recall curves and quantitative indexes are shown in Fig. 6. In order to reduce the influence of poor depth map in stereo saliency detection, depth confidence measure λ_d is introduced in our letter. Comparing the black line with the blue line in Fig. 6(a), it demonstrated that the performance with λ_d is superior to not using depth confidence measure, and the same conclusion can be drawn from the comparison of the first two columns in Fig. 6(b). At the stage of foreground saliency detection, a novel foreground seeds' selection mechanism using depth information, namely, DRSS, is proposed to acquire more accurate foreground seeds. As shown in Fig. 6, the DRSS process can achieve higher precision rate, and improve the performance of the final result according to the precision–recall curve.

IV. Conclusion

In this letter, a novel saliency detection model for stereoscopic images based on depth confidence analysis and multiple cues fusion is presented. First, the quality of depth map is considered when introduces the depth information into the saliency model, and a depth confidence measure is proposed to evaluate the reliability of depth map. In addition, a novel model for compactness integrating color and depth information is proposed to compute the compactness saliency. To achieve more robust saliency detection result, a foreground saliency detection method based on multiple cues contrast is proposed, which includes a novel DRSS method. Finally, weighted-sum method is used to generate the final saliency map. Experimental evaluation on two public benchmarks has validated the advantages of our approach.

REFERENCES

[1] Z. Ren, S. Gao, L. T. Chia, and I. Tsang, "Region-based saliency detection and its application in object recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 5, pp. 769–779, May 2014.

[2] Z. Wang, K. Liao, J. Xiong, and Q. Zhang, "Moving object detection based on temporal information," *IEEE Signal Process. Lett.*, vol. 21, no. 11, pp. 1403–1407, Nov. 2014.

[3] Y. Gao, M. Shi, D. Tao, and C. Xu, "Database saliency for fast image retrieval," *IEEE Trans. Multimedia*, vol. 17, no. 3, pp. 359–369, Mar. 2015.

[4] S. Han and N. Vasconcelos, "Image compression using object-based regions of interest," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2006, pp. 3097–3100.

[5] Y. Hou, P. Wang, W. Xiang, Z. Gao, and C. Hou, "A novel rate control algorithm for video coding based on fuzzy-PID controller," *Signal Image Video Process.*, vol. 9, no. 4, pp. 875–884, 2015.

[6] J. W. Yoo, S. Yea, and I. K. Park, "Content-driven retargeting of stereoscopic images," *IEEE Signal Process. Lett.*, vol. 20, no. 5, pp. 519–522, May 2013.

[7] R. Gallea, E. Ardizzone, and R. Pirrone, "Physical metaphor for streaming media retargeting," *IEEE Trans. Multimedia*, vol. 16, no. 4, pp. 971–979, Jun. 2014.

[8] Y. Fang, Z. Wang, W. Lin, and Z. Fang, "Video saliency incorporating spatiotemporal cues and uncertainty weighting," *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 3910–3921, Sep. 2014.

[9] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[10] M. M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu, "Global contrast based salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog. (CVPR)*, 2011, pp. 409–416.

[11] M.-M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr, and S.-M. Hu, "Global contrast based salient region detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 569–582, Mar. 2015.

[12] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang, "Saliency detection via graph-based manifold ranking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog. (CVPR)*, 2013, pp. 3166–3173.

[13] X. Wei, Z. Tao, C. Zhang, and X. Cao, "Structured saliency fusion based on Dempster–Shafer theory," *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1345–1349, Sep. 2015.

[14] J. Sun, H. Lu, and X. Liu, "Saliency region detection based on Markov absorption probabilities," *IEEE Trans. Image Process.*, vol. 24, no. 5, pp. 1639–1649, May 2015.

[15] L. Zhou, Z. Yang, Q. Yuan, Z. Zhou, and D. Hu, "Salient region detection via integrating diffusion-based compactness and local contrast," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3308–3320, Nov. 2015.

[16] H. Li, H. Lu, Z. Lin, X. Shen, and B. Price, "Inner and inter label propagation: Salient object detection in the wild," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3176–3186, Oct. 2015.

[17] H. Fu, D. Xu, S. Lin, and J. Liu, "Object-based RGBD image co-segmentation with mutex constraint," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog. (CVPR)*, 2015, pp. 4428–4436.

[18] J. Lei, J. Liu, H. Zhang, Z. Gu, N. Ling, and C. Hou, "Motion and structure information based adaptive weighted depth video estimation," *IEEE Trans. Broadcast.*, vol. 61, no. 3, pp. 416–424, Sep. 2015.

[19] Y. Fang, J. Wang, M. Narwaria, P. Le Callet, and W. Lin, "Saliency detection for stereoscopic images," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2625–2636, Jun. 2014.

[20] J. Lei, H. Zhang, L. You, C. Hou, and L. Wang, "Evaluation and modeling of depth feature incorporated visual attention for salient object segmentation," *Neurocomputing*, vol. 120, pp. 24–33, 2013.

[21] Y. Niu, Y. Geng, X. Li, and F. Liu, "Leveraging stereopsis for saliency analysis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog. (CVPR)*, 2012, pp. 454–461.

[22] K. Desingh, K. M. Krishna, D. Rajan, and C. V. Jawahar, "Depth really matters: Improving visual salient region detection with depth," in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, 2013, pp. 1–11.

[23] F. Guo, J. Shen, and X. Li, "Learning to detect stereo saliency," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, 2014, pp. 1–6.

[24] R. Ju, L. Ge, W. Geng, T. Ren, and G. Wu, "Depth saliency based on anisotropic center-surround difference," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2014, pp. 1115–1119.

[25] R. Ju, Y. Liu, T. Ren, L. Ge, and G. Wu, "Depth-aware salient object detection using anisotropic center-surround difference," *Signal Process. Image Commun.*, vol. 38, no. 10, pp. 115–126, 2015.

[26] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[27] D. Zhou, J. Weston, A. Gretton, O. Bousquet, and B. Schölkopf, "Ranking on data manifolds," in *Proc. Adv. Neural Inf. Process. Syst.*, 2004, pp. 169–176.

[28] B. Alexe, T. Deselaers, and V. Ferrari, "Measuring the objectness of image windows," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2189–2202, Nov. 2012.

[29] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 17, pp. 971–987, Jul. 2002.

[30] Y. Fang, W. Lin, Z. Fang, Z. Chen, C.-W. Lin, and C. Deng, "Visual acuity inspired saliency detection by using sparse features," *Inf. Sci.*, vol. 309, pp. 1–10, 2015.

[31] Y. Qin, H. Lu, Y. Xu, and H. Wang, "Saliency detection via cellular automata," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog. (CVPR)*, 2015, pp. 110–119.

[32] C. Li, Y. Yuan, W. Cai, Y. Xia, and D. D. Feng, "Robust saliency detection via regularized random walks ranking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog. (CVPR)*, 2015, pp. 2710–2717.