Co-saliency Detection for RGBD Images Based on Multi-constraint Superpixels Matching and Co-cellular Automata

Zhengyi Liu[1(✉)] and Feng Xie[2]

[1] Key Laboratory of Intelligent Computing and Signal Processing,
Ministry of Education, Anhui University, Hefei, China
`22927463@qq.com`
[2] Co-Innovation Center for Information Supply and Assurance Technology,
Hefei, China
`811173098@qq.com`

Abstract. Co-saliency detection aims at extracting the common salient regions from an image group containing two or more relevant images. It is a newly emerging topic in computer vision community. Different from the existing co-saliency methods focusing on RGB images, this paper proposes a novel co-saliency detection model for RGBD images, which utilizes the depth information to enhance identification of co-saliency. First, we utilize the existing single saliency maps as the initialization, then we use multiple cues to compute combination inter-images similarity to match inter-neighbors for each superpixel. Especially, we extract high dimensional features for each image region with a deep convolutional neural network as semantic cue. Finally, we introduce a modified 2-layer Co-cellular Automata to exploit depth information and the intrinsic relevance of similar regions through interactions with neighbors in multi-scene. The experiments on two RGBD co-saliency datasets demonstrate the effectiveness of our proposed framework.

Keywords: RGBD · Co-saliency · Cellular automata · Semantic feature
Multi-constraint

1 Introduction

In recent years, co-saliency detection has become an emerging issue in saliency detection, which detects the common salient regions among multiple images [1–4]. Different from the traditional single saliency detection model, co-saliency detection model aims at discovering the common salient objects from an image group containing two or more relevant images, while the categories, intrinsic characteristics, and locations of the salient objects are entirely unknown [5]. The co-salient objects simultaneously exhibit two properties, i.e. (1) The co-salient regions should be salient with respect to the background in each image, and (2) All these co-salient regions should be similar in appearance among multiple images. Due to its superior expansibility, co-saliency detection has been widely used in many computer vision tasks, such as foreground co-segmentation [6], object co-localization and detection [7], and image matching [8].

Most existing co-saliency detection models are focused on RGB images and have achieved satisfactory performances [9–16]. Recently, Co-saliency detection for RGBD images has become one of the popular and challenging problem. RGBD co-saliency detection in [17] is firstly discussed. They proposed a RGBD co-saliency model using bagging-based clustering. Then, Cong et al. [18] proposed an iterative RGBD co-saliency framework, which utilized the existing single saliency maps as the initialization, and generated the final RGBD co-saliency map by using a refinement-cycle model. In their another paper [19], they proposed a co-saliency model based on multi-constraint feature matching and cross label propagation. In this paper, for combining depth and repeatability, we firstly propose a matching algorithm based on neighboring superpixel sets of Multi-Constraint distance to calculate the similarity between images and to depict the occurrence of area repetition. Secondly, inspired by Ref. [23], we propose a 2-Layer co-cellular automata model to calculate the saliency spread of intra-images and inter-images, in order to ensure complete saliency of targeted area. Besides, the depth information and high dimensional features are considered in our method to achieve better result. The major contributions of the proposed co-saliency detection method are summarized as follows.

(1) We extract high dimensional features for each image region with a deep convolutional neural network as semantic cue and combine it with color cue, depth cue, and saliency cue to calculate the similarity between two superpixels for the first time.
(2) A modified 2-layer co-cellular automata model is used to calculate the saliency spread of intra-images and inter-images, in order to ensure complete saliency of targeted area.
(3) Both semantic information and depth information are considered in cellular automata to optimize this co-saliency model in our method.

The rest of this paper is organized as follows. Section 2 introduces the proposed method in detail. The experimental results with qualitative and quantitative evaluations are presented in Sect. 3. Finally, the conclusion is drawn in Sect. 4.

2   Proposed Method

The proposed RGBD co-saliency framework is introduced in this section. Figure 1 shows the framework of the proposed method. Our method is initialized by the existing single saliency maps, and then we propose a matching algorithm based on neighboring superpixel sets of Multi-Constraint distance to calculate the similarity between images and to depict the occurrence of area repetition. Finally, inspired by Ref. [23], we propose a 2-Layer co-cellular automata model to calculate the saliency spread of intra-images and inter-images, in order to ensure complete saliency of targeted area.

Notations: Given N input images $\{I^i\}_{i=1}^N$, and the corresponding depth maps are denoted as $\{D^i\}_{i=1}^N$. The M_i single saliency maps for image I^i produced by existing single image saliency models are represented as $S^i = \left\{S_j^i\right\}_{j=1}^{M_i}$. In our method, the superpixel-level region is regarded as the basic unit for processing. Thus, each RGB image I^i is abstracted into superpixels $R^i = \left\{r_m^i\right\}_{m=1}^{N_i}$ using SLIC algorithm [24] firstly, where N_i is the number of superpixels for image I^i.
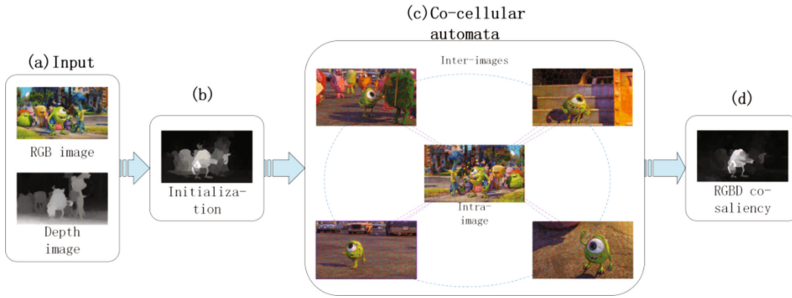

Fig. 1. The framework of our algorithm. (a) Input RGB image and the corresponding depth map. (b) Initialization. (c) Superpixel matching and parallel evolution via co-cellular automata. (d) The final saliency result.

2.1 Initialization

The proposed co-saliency framework aims at discovering the co-salient objects from multiple images in a group with the assistance of existing single saliency maps. Therefore, some existing saliency maps produced by single saliency models are used to initialize the framework. It is well known that different saliency methods own different superiority in detecting salient regions. In a way, these saliency maps are complementary in some regions, thus, the fused result can inherit the merits of the multiple saliency maps, and produce more robust and superior detection baseline. In our method, the simple average function is used to achieve a more generalized initialization result. The initialized saliency map for image I^i is denoted as:

$$S_f^i\left(r_m^i\right) = \frac{1}{M}\sum_{j=1}^{M_i} S_j^i\left(r_m^i\right) \tag{1}$$

Where $S_j^i\left(r_m^i\right)$ denotes the saliency value of superpixel r_m^i produced by j^{th} saliency method for image I^i. In our experiments, four saliency methods including RC [20], DCLC [21], RRWR [22], and BSCA [23], are used to produce the initialized saliency map.

2.2   Superpixel Matching via Multi-constraint Cues

For convenience of calculations and intrinsic structural information, the image is firstly segmented into a set of superpixels by simple linear iterative clustering (SLIC) algorithm [24]. The core of detecting the common salient object is the superpixel matching in different images. In this paper, superpixel matching means, for any superpixel r_m^i in image I^i, finding a set of superpixels with high similarity in another image I^j. Note that not all superpixels can be matched and one superpixel can have several matching superpixels in other images. In this paper, high-dimensional semantic cue and low-dimensional cue are both utilizing to compute the similarity between images.

High-Dimensional Cue. We extract high-dimensional features for each image region with a deep convolutional neural network originally trained over the ImageNet dataset using Caffe, an open source framework for CNN training and testing. The architecture of this CNN has eight layers including five convolutional layers and three fully-connected layers. Features are extracted from the output of the second last fully connected layer, which has 4096 neurons. Although this CNN was originally trained on a dataset for visual recognition, automatically extracted CNN features turn out to be highly versatile and can be more effective than traditional handcrafted features on other visual computing tasks.

Since an image region may have an irregular shape while CNN features have to be extracted from a rectangular region, to make the CNN features only relevant to the pixels inside the region, we define the rectangular region for CNN feature extraction to be the bounding box of the image region and fill the pixels outside the region but still inside its bounding box with the mean pixel values at the same locations across all ImageNet training images. These pixel values become zero after mean subtraction and do not have any impact on subsequent results. We warp the region in the bounding box to a square with 227×227 pixels to make it compatible with the deep CNN trained for ImageNet. The warped RGB image region is then fed to the deep CNN and a 4096-dimensional feature vector is obtained by forward propagating a mean-subtracted input image region through all the convolutional layers and fully connected layers. We name this vector feature F.

Thus, the high-dimensional semantic similarity is defined as:

$$S_h\left(r_m^i, r_n^j\right) = \exp\left(-\frac{\left\|F_m^i - F_k^j\right\|_2}{\sigma^2}\right) \qquad (2)$$

where F_m^i denotes 4096 high-dimensional features contrast of superpixel r_m^i, and σ^2 is a constant.

Low-Dimensional Cue. Three low-dimensional cues include color cue, depth cue, and saliency cue are used to gain a multi-constraint cue.

RGB Similarity. The color histogram [25] are used to represent the RGB feature on the superpixel level, which are denoted as HC_m^i. Then, the Chi-square measure is employed to compute the feature difference. Thus, the RGB similarity is defined as:

$$S_c\left(r_m^i, r_n^j\right) = 1 - \frac{1}{2}\chi^2\left(HC_m^i, HC_n^j\right) \tag{3}$$

where r_m^i and r_n^j are the superpixels in image I^i and I^j, respectively, and $\chi^2(\cdot)$ denotes the Chi-square distance function.

Depth Similarity. Two depth consistency measurements, namely depth value consistency and depth contrast consistency, are composed of the final depth similarity measurement, which is defined as:

$$S_d\left(r_m^i, r_n^j\right) = \exp\left(-\frac{W_d\left(r_m^i, r_n^j\right) + W_c\left(r_m^i, r_n^j\right)}{\sigma^2}\right) \tag{4}$$

where $W_d\left(r_m^i, r_n^j\right)$ is the depth value consistency measurement to evaluate the inter-image depth consistency, due to the fact that the common regions should appear similar depth values.

$$W_d\left(r_m^i, r_n^j\right) = \left|d_m^i - d_n^j\right| \tag{5}$$

$W_c\left(r_m^i, r_n^j\right)$ describes the depth contrast consistency, because the common regions should represent more similar characteristic in depth contrast measurement.

$$W_c\left(r_m^i, r_n^j\right) = \left|D_c\left(r_m^i\right) - D_c\left(r_n^j\right)\right| \tag{6}$$

with

$$D_c\left(r_m^i\right) = \sum_{k \neq m}\left|d_m^i - d_n^j\right|\exp\left(-\frac{\left\|p_m^i - p_k^i\right\|_2}{\sigma^2}\right) \tag{7}$$

where $D_c\left(r_m^i\right)$ denotes the depth contrast of superpixel r_m^i, p_m^i denotes the position of superpixel r_m^i, and σ^2 is a constant.

Saliency Similarity. Inspired by the prior that the common regions should appear more similar in single saliency map compared to other regions, the output saliency map from the addition scheme is used to define the saliency similarity measurement in our work:

$$S_s\left(r_m^i, r_n^j\right) = \exp\left(-\left|S_{sp}^i\left(r_m^i\right) - S_{sp}^j\left(r_n^j\right)\right|\right) \tag{8}$$

where $S_{sp}^i\left(r_m^i\right)$ is saliency score of superpixel r_m^i via initialization.

Based on these cues, the combination similarity measurement is defined as the average of the four similarity measurements.

$$S_M\left(r_m^i, r_n^j\right) = \frac{S_h\left(r_m^i, r_n^j\right) + S_c\left(r_m^i, r_n^j\right) + S_d\left(r_m^i, r_n^j\right) + S_s\left(r_m^i, r_n^j\right)}{4} \tag{9}$$

where $S_h\left(r_m^i, r_n^j\right)$, $S_c\left(r_m^i, r_n^j\right)$, $S_d\left(r_m^i, r_n^j\right)$, and $S_s\left(r_m^i, r_n^j\right)$ are the normalized semantic, RGB, depth, and saliency similarities between superpixel r_m^i and r_n^j, respectively. A larger $S_M\left(r_m^i, r_n^j\right)$ value corresponds to greater similarity between two superpixels.

2.3   Co-saliency Detection via 2-Layer Co-cellular Automata

In Ref. [23], Cellular Automata method was proposed to calculate the saliency of a single image. The core concept of this method is that the saliency of one superpixel is affected by itself and the adjacent superpixels. All of the superpixels will converge after several times of spread. However, for co-saliency detection, as shown in Fig. 2, the saliency of one superpixel is affected by its intra-neighbor (blue and yellow spots) and its inter-neighbor (purple spot) at the same time.

According to this theory, we propose 2-layer Co-cellular Automata via intra image and inter images spread:

$$S_{m+1}^i = (1 - \kappa_1 - \kappa_2)S_m^i + \kappa_1 F_{\text{intra}}^i S_m^i + \kappa_2 \sum_{j=1, j \neq i}^{n} F_{\text{inter}}^{i,j} S_m^j \tag{10}$$

where S_m^i is the saliency of all superpixels in I^i after m times of status updates, S_0^i is the initial saliency via Eq. (1), F_{intra}^i is the influence matrix of superpixels in I^i, $F_{\text{inter}}^{i,j}$ is the influence matrix from I_j to I_i, κ_1 and κ_2 are impact factors. In this model, we utilize the structural information of intra-image, also, the corresponding relationship is considered here.
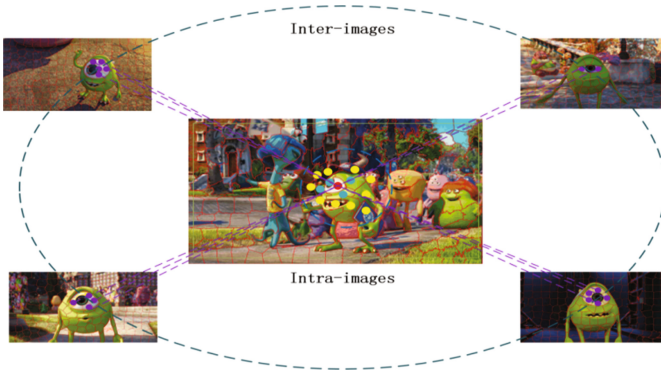

Fig. 2. Co-saliency detection model. The saliency of one superpixel (red spots) is not only affected by the adjacent superpixels (blue and yellow spots) but also affected by the matched superpixels in other images (purple spots). (Color figure online)

Intra-image Influence Matrix. In Ref. [23], the similarity of intra-image superpixels is calculated by color similarity in CIELab color space. Here, we also consider the affect of depth cue and semantic cue. We define the initial intra-image influence matrix as $F'^i_{intra} = \left[f^i_{s,t} \right]_{N^i \times N^i}$.

$$f^i_{s,t} = \begin{cases} \exp\left(-\dfrac{\left\| c^i_s - c^i_t \right\|^2 + \left\| d^i_s - d^i_t \right\| + \left\| F^i_s - F^i_t \right\|^2}{2\sigma_f^2} \right) & t \in N^i_s \\ 0 & t = s \text{ or others} \end{cases} \tag{11}$$

Where N^i_s is superpixels's 2-layer adjacent region (not only includes its neighbor, but also its neighbor's neighbor). In order to normalize impact factor matrix, a degree matrix $D^i_{intra} = diag\{d_1, d_2, \ldots, d_{N^i}\}$, where $d_i = \sum_t f^i_{s,t}$. Finally, a row-normalized impact factor matrix can be clearly calculated as follows:

$$F^i_{intra} = \left[D^i_{intra} \right]^{-1} \bullet F'^i_{intra} \tag{12}$$

Inter-Image Influence Matrix. To utilize the affect of other images in the same set, we use the method introduced in Sect. 2.2 to obtain $S_M\left(r^i_m, r^j_n \right)$, then the initial inter-image influence matrix is defined as $F'^i_{inter}\left[f^{i,j}_{s,t} \right]_{N^i * N^j}$ to capture the relationship of any two superpixels in different images.

$$f^{i,j}_{s,t} = \begin{cases} S_M\left(r^i_m, r^j_n \right) & S_M\left(r^i_m, r^j_n \right) > \delta \\ 0 & \text{others} \end{cases} \tag{13}$$

where δ is a threshold to match saliency. Here this parameter is set to be 0.9 according to our experience. Same as above, degree matrix $D^i_{inter} = diag\{d_1, d_2, \ldots d_{N^i}\}$, where $d_i = \sum_t f^{i,j}_{s,t}$. And the row-normalized impact factor matrix is indicated as:

$$F^i_{inter} = \frac{1}{N-1} \bullet \left[D^i_{inter} \right]^{-1} \bullet F'^i_{inter} \tag{14}$$

The overall framework of the proposed method is summarized in Table 1.

Table 1. The procedure of our method.

Alogrithm 1. The Overall Framework.
Input: The RGB images and depth maps in an image group.
Output: The co-saliency map for each image.
1: for each image in the group do
2: Obtain the initialized saliency map using Eq. (1) ;
3: end for
4: for each image in the group do
5: Calculate the intra-image impact factor matrix $F_{\mathrm{int}ra}^{i}$ using Eq. (11-12);
6:   Calculate matching similarity $S_{M}\left(r_{m}^{i}, r_{n}^{j}\right)$ of any two superpixels in different images using Eq. (2-9);
7: Obtain the inter-image impact factor matrix $F_{\mathrm{int}er}^{i}$ using Eq. (13-14);
8: while m<M do
9: update saliency map using Eq. (10);
10: end while
11: end for

3   Experiment

In this section, we would evaluate the proposed RGBD co-saliency framework on two RGBD co-saliency datasets. The qualitative and quantitative comparison with other state-of-the-art methods are presented.

3.1   Experimental Settings

Two RGBD benchmarks: the RGBD Coseg183 dataset [27] and the RGBD Cosal150 dataset [18] are used to evaluate our method. The RGBD Coseg183 dataset is composed of 183 pictures, and these pictures are distributed in 16 groups. And the RGBD Cosal150 dataset contains 150 images that are distributed in 21 image sets.

We adopted two quantitative criteria to evaluate the co-saliency map, which is Precision-Recall(PR) curve and F-measure score. The precision and recall score are computed by ground truth. F-measure [28] is defined as the weighted mean of precision P and recall R, which is denoted as:

$$F - measure = \frac{\left(1 + \beta^2\right) \times P \times R}{\beta^2 \times P + R} \tag{15}$$

Where β^2 is set to 0.3, because the precision is more important than Recall.

In this method, the number of superpixels of each image is set to 200, the maximum number of iterations M is set to 20. And the parameter κ_1 and κ_2 in Eq. (9) is set to 0.3 and 0.5, respectively.

3.2   Comparison with State-of-the-Art Methods

In this section, we compare our method with 10 state-of-the-art methods, which are RC [20], DCLC [21], RRWR [22], BSCA [23], SE [33], FP [34], CCS [4], EMR [13], AIF [18] and MCLP [19]. The first four methods are single image saliency methods, also they are regarded as the input of this method. SE and FP are classic RGBD single saliency algorithms. CCS and EMR are co-saliency methods for RGB images. The last two method are the latest co-saliency methods for RGDB images.
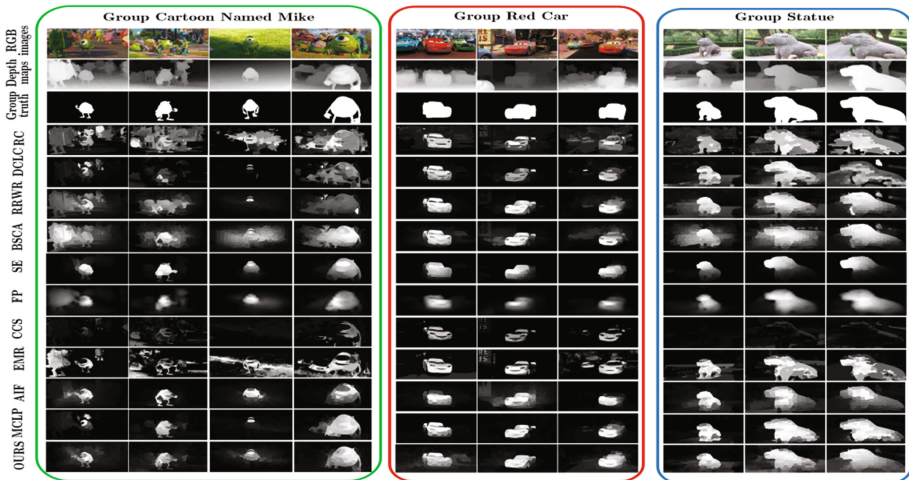

Fig. 3. Visual comparison of different saliency and co-saliency detection methods on two datasets.

Some visual examples on two datasets are shown in Fig. 3. The quantitative comparison results including the PR curves and F-measure scores are reported in Fig. 4. As can be seen, on the RGBD Cosal150 dataset, the proposed method's curve intersects with SE, FP, AIF and MCLP, but the F-measure score of the proposed method is only

slightly lower than MCLP. In contrast, the RGBD Coseg183 dataset is more difficult and challenging for co-saliency detection, however, the proposed method's curve achieves the highest precision of the whole PR curves, and the F-measure is only slightly lower than MCLP, too.
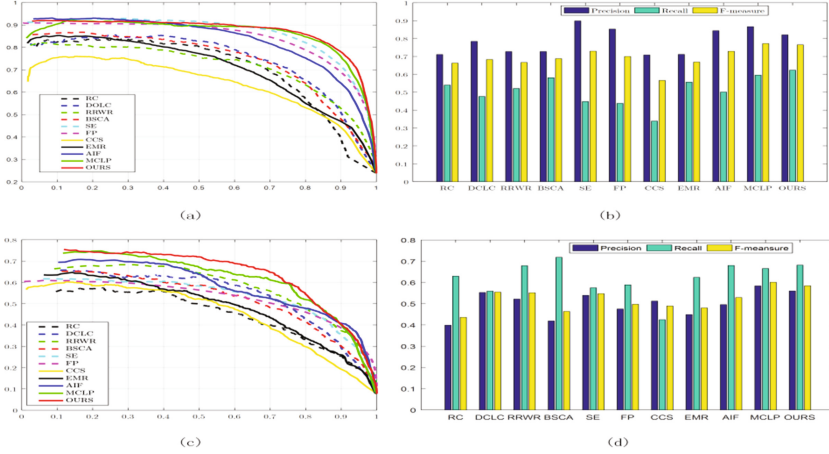

Fig. 4. Quantitative comparisons between the proposed method and the state-of-the-art methods on two datasets. "OURS" means the proposed method. (a) and (b) are PR curves and F-measure scores on RGBD Cosal150 dataset, respectively. (c) and (d) are PR curves and F-measure scores on RGBD Coseg183 dataset, respectively.

Table 2 shows the comparison of average run time to process one image between our proposed model and the other two RGBD co-saliency detection methods (AIF, MCLP). The measurement environment is Intel (R) Core (TM) i5-4570 CPU 3.20.GHz workstation with 8 GB RAM under Matlab R2012a platform. It can be seen from the table that our proposed algorithm is faster than AIF and MCLP.

Table 2. Comparison of average run time to process one image between different algorithm.

Method	Run time (s)
AIF	42.67
MCLP	41.03
OURS	13.07

4   Conclusion

In this paper, we present a co-saliency detection model for RGBD images, which utilize mutli-constraint cues to capture the relationship among multiple images for superpixels matching. Further on, impact factor matrix are constructed for intra-images and inter-images, and the depth cue and high-dimensional semantic cue are considered in intra-images impact factor matrix constructing. In the end, a modified 2-layer co-cellular

automata model is using to update initial saliency maps. The comprehensive comparison and discussion on two RGBD co-saliency datasets have demonstrated that the proposed method outperforms other state-of-the-art saliency and co-saliency models.

References

1. Chen, H.T.: Preattentive co-saliency detection. In: IEEE International Conference on Image Processing 2010, vol. 119, pp. 1117–1120. IEEE (2010)
2. Li, H., Ngan, K.N.: A co-saliency model of image pairs. IEEE Trans. Image Process. **20**(12), 3365–3375 (2011)
3. Chang, K.Y., Liu, T.L., Lai, S.H.: From co-saliency to co-segmentation: an efficient and fully unsupervised energy minimization model. In: IEEE Conference on Computer Vision and Pattern Recognition 2011, vol. 42, pp. 2129–2136. IEEE Computer Society (2011)
4. Fu, H., Cao, X., Tu, Z.: Cluster-based co-saliency detection. IEEE Trans. Image Process. **22** (10), 3766–3778 (2013)
5. Zhang, D., Fu, H., Han, J., Borji, A., Li, X.: A review of co-saliency detection technique: fundamentals, applications, and challenges (2016)
6. Fu, H., Xu, D., Zhang, B., Lin, S., Ward, R.K.: Object-based multiple foreground video co-segmentation via multi-state selection graph. IEEE Trans. Image Process. **24**(11), 3415–3424 (2015)
7. Tang, K., Joulin, A., Li, L.J., Li, F.F.: Co-localization in real-world images. In: IEEE Computer Vision and Pattern Recognition 2014, pp. 1464–1471. IEEE (2014)
8. Toshev, A., Shi, J., Daniilidis, K.: Image matching via saliency region correspondences. In: IEEE Computer Vision and Pattern Recognition 2007, pp. 1–8. IEEE (2007)
9. Liu, Z., Zou, W., Li, L., Shen, L., Meur, O.L.: Co-saliency detection based on hierarchical segmentation. IEEE Sig. Process. Lett. **21**(1), 88–92 (2013)
10. Ge, C., Fu, K., Liu, F., Bai, L., Yang, J.: Co-saliency detection via inter and intra saliency propagation. Sig. Process. Image Commun. **44**(C), 69–83 (2016)
11. Cao, X., Tao, Z., Zhang, B., Fu, H., Li, X.: Saliency map fusion based on rank-one constraint. In: IEEE International Conference on Multimedia and Expo 2013, pp. 1–6 (2013)
12. Cao, X., Tao, Z., Zhang, B., Fu, H., Feng, W.: Self-adaptively weighted co-saliency detection via rank constraint. IEEE Trans. Image Process. Publ. IEEE Signal Process. Soc. **23**(9), 4175–4186 (2014)
13. Li, Y., Fu, K., Liu, Z., Yang, J.: Efficient saliency-model-guided visual co-saliency detection. IEEE Signal Process. Lett. **22**(5), 588–592 (2014)
14. Huang, R., Feng, W., Sun, J.: Saliency and co-saliency detection by low-rank multiscale fusion. In: IEEE International Conference on Multimedia and Expo 2015, pp. 1–6 (2015)
15. Zhang, D., Han, J., Li, C., Wang, J.: Co-saliency detection via looking deep and wide. In: IEEE Computer Vision and Pattern Recognition 2015, pp. 2994–3002 (2015)
16. Zhang, D., Meng, D., Li, C., Jiang, L.: A self-paced multiple-instance learning framework for co-saliency detection. In: IEEE International Conference on Computer Vision 2015, pp. 594–602 (2015)
17. Song, H., Liu, Z., Xie, Y., Wu, L., Huang, M.: RGBD co-saliency detection via bagging-based clustering. IEEE Sig. Process. Lett. **23**(12), 1722–1726 (2016)
18. Cong, R., Lei, J., Fu, H., Lin, W., Huang, Q., Cao, X., et al.: An iterative co-saliency framework for RGBD images. IEEE Trans. Cybern. **PP**(99), 1–14 (2017)
19. Cong, R., et al.: Co-saliency detection for RGBD images based on multi-constraint feature matching and cross label propagation. IEEE Trans. Image Process. **PP**(99), 1 (2018)

20. Cheng, M.M., Zhang, G.X., Mitra, N.J., Huang, X., Hu, S.M.: Global contrast based salient region detection. In: IEEE CVPR 2011, vol. 37, pp. 409–416 (2011)
21. Zhou, L., Yang, Z., Yuan, Q., Zhou, Z., Hu, D.: Salient region detection via integrating diffusion-based compactness and local contrast. IEEE Trans. Image Process. **24**(11), 3308–3320 (2015)
22. Li, C., Yuan, Y., Cai, W., Xia, Y.: Robust saliency detection via regularized random walks ranking, pp. 2710–2717 (2015)
23. Qin, Y., Lu, H., Xu, Y., Wang, H.: Saliency detection via cellular automata. In: IEEE Computer Vision and Pattern Recognition 2015, pp. 110–119 (2015)
24. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: Slic superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Anal. Mach. Intell. **34**(11), 2274–2282 (2012)
25. Leung, T., Malik, J.: Recognizing surfaces using three-dimensional textons. In: IEEE International Conference on Computer Vision 1999, vol. 2, pp. 1010–1017 (1999)
26. Fu, H., Xu, D., Lin, S., Liu, J.: Object-based RGBD image co-segmentation with mutex constraint. In: IEEE Conference on Computer Vision and Pattern Recognition 2017, pp. 4428–4436 (2017)
27. Achanta, R., Hemami, S., Estrada, F., Susstrunk, S.: Frequency-tuned salient region detection. In: Proceedings of CVPR, June 2009, pp. 1597–1604 (2009)
28. Ju, R., Liu, Y., Ren, T., Ge, L., Wu, G.: Depth-aware salient object detection using anisotropic center-surround difference. Sig. Process. Image Commun. **38**(C), 115–126 (2015)
29. Peng, H., Li, B., Xiong, W., Hu, W., Ji, R.: RGBD salient object detection: a benchmark and algorithms. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8691, pp. 92–109. Springer, Cham (2014) https://doi.org/10.1007/978-3-319-10578-9_7
30. Feng, D., Barnes, N., You, S., Mccarthy, C.: Local background enclosure for RGB-D salient object detection. In: IEEE Conference on Computer Vision and Pattern Recognition 2016, pp. 2343–2350. IEEE Computer Society (2016)
31. Cong, R., Lei, J., Zhang, C., Huang, Q., Cao, X., Hou, C.: Saliency detection for stereoscopic images based on depth confidence analysis and multiple cues fusion. IEEE Sig. Process. Lett. **23**(6), 819–823 (2016)
32. Quo, J., Ren, T., Bei, J.: Salient object detection for RGB-D image via saliency evolution. In: IEEE International Conference on Multimedia and Expo, pp. 1–6. IEEE (2016)
33. Guo, J., Ren, T., Bei, J.: Salient object detection in RGB-D image based on saliency fusion and propagation. In: 7th International Conference on Internet Multimedia Computing and Service, 59. ACM (2015)
34. Li, H., Lu, H., Lin, Z., Shen, X., Price, B.: Inner and inter label propagation: salient object detection in the wild. The first five years of the Communist International. New Park Pub., 3176–3186 (1973)
35. Chen, T., Lin, L., Liu, L., Luo, X., Li, X.: DISC: deep image saliency computing via progressive representation learning. IEEE Trans. Neural Netw. Learn. Syst. **27**(6), 1135 (2015)
36. He, S., Lau, R.W.H., Liu, W., Huang, Z., Yang, Q.: Supercnn: a superpixelwise convolutional neural network for salient object detection. Int. J. Comput. Vision **115**(3), 330–344 (2015)
37. Lee, G., Tai, Y.W., Kim, J.: Deep saliency with encoded low level distance map and high level features. In: Computer Vision and Pattern Recognition, pp. 660–668. IEEE (2016)
38. Li, G., Yu, Y.: Deep contrast learning for salient object detection. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 478–487. IEEE Computer Society (2016)