

# RGB-D Saliency Detection with Multi-feature-fused Optimization

Tianyi Zhang<sup>1</sup>, Zhong Yang<sup>1,2(✉)</sup>, and Jiarong Song<sup>1</sup>

<sup>1</sup> College of Automation Engineering, Nanjing University of Aeronautics and Astronautics,  
Nanjing 210016, People's Republic of China  
YangZhong@nuaa.edu.cn

<sup>2</sup> Unmanned System Laboratory, Nanjing University of Aeronautics and Astronautics,  
Nanjing 210016, People's Republic of China

**Abstract.** This paper proposes a three-stage method using color, joint entropy and depth to detect salient regions in an image. In the first stage, coarse saliency maps are computed through multi-feature-fused manifold ranking. However, graph-based saliency detection methods like manifold ranking often have problems of inconsistency and misdetection if some parts of the background or objects have relatively high contrast with its surrounding areas. To solve this problem and provide more robust results in varying conditions, depth information is repeatedly used to segment and refine saliency maps. In details, a self-adaptive segment method based on depth distribution is used secondly to filter the less significant areas and enhance the salient objects. At last, the saliency-depth consistency check is implemented to suppress the highlighted areas in the background and enhance the suppressed parts of objects. Qualitative and quantitative evaluation on a challenging RGB-D dataset demonstrates significant appeal and advantages of our algorithm compared with eight state-of-the-art methods.

**Keywords:** RGB-D saliency detection · Multi-feature fusion  
Manifold ranking · Saliency-depth consistency check

## 1 Introduction

Salient object detection is a subject aiming to detect the most salient attention-attracted object in a scene [1]. As an important part of computer vision and neuroscience, its application has been widely expand to various fields [2] such as image segmentation [3], retargeting [4, 5], compressing [6], object recognition [7] and object class discovery [8, 9]. Over the past decades, a lot of algorithms have been proposed for detecting salient regions in images, most of which aim to identify the salient subset from images first and then use it as a basis to segment salient objects from the whole image [10]. Generally, these approaches can be categorized into two breakdowns: bottom-up and top-down approaches. Bottom-up saliency approaches use some pre-assumed priors like contrast prior and texture prior to compute saliency of image while top-down approaches usually use high-level information to guide the detection. In this paper, only bottom-up saliency detection methods are considered.

To our best knowledge, bottom-up saliency detection methods can be categorized as patch-based saliency detection methods and graph-based saliency methods. Patch-based saliency detection utilizes pixels and patches as the basic components for saliency calculation while graph-based saliency detection regards each superpixel segmented from image as a node in graph and calculates saliency according to graph theory. Compared with pixel-based methods, graph-based saliency detection methods can more easily and directly represent the spatial relationship between superpixels.

Among all studies in patch-based saliency detection, Chen *et al.* [11] provided a simple but power approach to compute the saliency of each sparse-represented color. Duan *et al.* [12] presented a saliency detection method based on the spatially weighted dissimilarity which integrated dissimilarities, dimensional space and the central bias to determine the contrast of a patch. Other methods on pixel-based saliency detection also have made great contribution to detect attention-attracted regions. However, their results are sensitive to small regions with strong contrast, especially on the edge of image. For graph-based approaches, Chuan Yang *et al.* [13] proposed a graph-based manifold ranking algorithm to measure the difference between four sides and the rest parts of image. Yao Qin *et al.* [14] constructed a background-based map using colour and space contrast with the clustered boundary seeds to initial saliency maps. Other graph-based methods also show an outstanding performance on boundary suppressing. However, in order to enhance the local contrast, most of graph-based methods only assign edges between superpixels and their adjacent nodes. Because of this, some regions could be ununiformed, which will significantly decrease the accuracy of detection.

Although the previous methods mentioned above showed good results in some scenes, there are still a lot of problems remaining to be solved. The Saliency regions of an image often have great contrast in color, texture or other features. However, some parts of background can also have high-contrast regions while the contrast of some salient parts can be relatively low. This phenomenon will cause inconsistency and misdetection in saliency maps and cannot be removed by neither graph-based nor pixel-based algorithm due to the nature of contrast comparison.

The depth information of object has an indispensable role in the vision ability of human beings according to cognitive psychology [15]. Recently, with the development of stereo vision, smoothed and accurate depth of an image can be easily acquired through Kinect, ZED camera or other stereo cameras. Because of this, there is an increasing number of methods [16–18], both pixel-based and graph-based, that introduce depth as an important basis for saliency detection. To take advantage of it, this paper repeatedly uses depth information as a basis to generate, filter and refine saliency maps. Firstly, a multi-feature-fused manifold ranking algorithm is suggested to generate coarse saliency maps. Compared with traditional manifold ranking [13], this algorithm is more robust to varying scenes by taking depth and texture as references for ranking. Then, a self-adaptive segmentation method based on depth distribution is used to filter the less salient areas. This step segments possible salient regions from candidates based on the depth information of each superpixel. At last, we propose a further method of saliency-depth consistency check to remove or enhance ununiformed regions in background and saliency regions respectively. In the first step of this algorithm, multiple features are taken as references for ranking, some of which are unnecessary and can cause error

highlighting or suppressing. The error saliency of these areas will be refined in the third step due to their saliency inconsistent with depth values.

## 2 Proposed Method

Our algorithm mainly includes three steps. Firstly, we generate coarse saliency maps through multi-feature-fused manifold ranking [13]. In this process, color, depth and joint entropy of superpixels are taken as references for ranking. Then, a self-adaptive segmentation method based on depth distribution is used to furtherly erase the saliency of background. Finally, we implement saliency-depth consistency check to remove small speckles on the background and enhance ununiformed regions on objects. The main stages of our algorithm are demonstrated in Fig. 1. In the figure, the result of every stage is listed to illustrate how our algorithm optimizes the saliency map step by step.

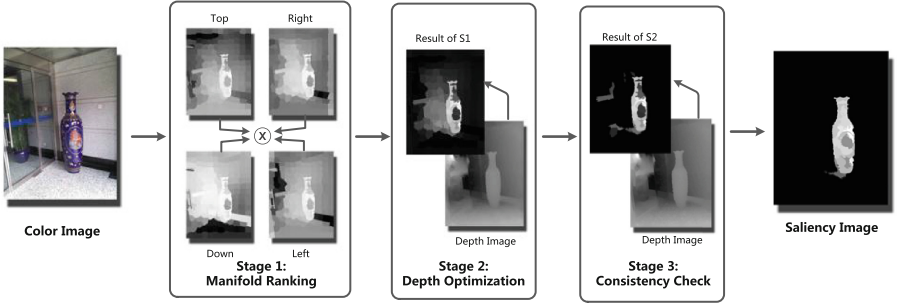


Fig. 1. Illustration of the main stages of proposed algorithm

### 2.1 Multi-feature-fused Manifold Ranking

Suppose  $X = \{x_1, \dots, x_i, x_{i+1}, \dots, x_n\} \in \mathbb{R}^{m \times n}$  is a data set, some of which are labelled as queries and others are supposed to be ranked according to their similarity of the queries. In our case,  $X$  is a set of superpixels segment via simple linear iterative clustering (SLIC) method [19].  $f: X \rightarrow \mathbb{R}^n$  is a ranking function which assigns each superpixel  $x_i$  to a ranking score  $r_i$ . Then, a graph  $G = (V, E)$  is constructed, where nodes  $V$  are superpixels in  $X$ .  $W \in \mathbb{R}^{n \times n}$  denotes the adjacency matrix with element  $w_{ij}$  to store the edges  $E$ , which is the weight between superpixel  $i$  and  $j$ . Finally,  $y = [y_1, \dots, y_n]^T$  is an initial vector in which  $y_i = 1$  if  $x_i$  is a query and  $y_i = 0$  otherwise.

The cost function associated with  $r_i$  is defined to be:

$$f^* = \arg \min_f \frac{1}{2} \left( \sum_{i,j=1}^n w_{ij} \left\| \frac{f_i}{\sqrt{d_{ii}}} - \frac{f_j}{\sqrt{d_{jj}}} \right\|^2 + \mu \sum_{i=1}^n \|f_i - y_i\|^2 \right) \quad (1)$$

The first and second term of cost function is smoothness constrain and fitting constrain, respectively. In this paper, a unnormalized Laplacian matrix is used to construct  $f^*$ :

$$f^* = (D - \frac{1}{\mu + 1}W)^{-1}y \quad (2)$$

where  $D = \text{diag}\{d_{11}, \dots, d_{nn}\}$  is the degree matrix of  $G$ ,  $d_{ii} = \sum_j w_{ij}$ .  $\mu$  is the parameter which controls the balance between smoothness constrain and fitting constrain.

Traditional manifold ranking algorithm just assigns the Euclidean distance of colour in CIElab space to adjacency matrix, which does not consider other features such as depth. In this paper, we define  $W$  as:

$$w_{ij} = \begin{cases} \exp(-\frac{1}{2\sigma_c^2}c_{ij} - \frac{1}{2\sigma_d^2}d_{ij} - \frac{1}{2\sigma_h^2}h_{ij}) & j \in N_i \\ 0 & otherwise \end{cases} \quad (3)$$

where  $N_i$  is the set of superpixels which are adjacent to  $i$ .  $\sigma_c, \sigma_d, \sigma_h$  are weighting factors for the corresponding features  $c_{ij}, d_{ij}, h_{ij}$ , which have been normalized to  $[0, 1]$  before fusion.  $c_{ij} = \|c_i - c_j\|$  is the Euclidean distance of colour in CIElab space.  $d_{ij} = \|d_i - d_j\|$  is the Euclidean distance of depth between two superpixels.  $h_{ij}$  refers to the joint entropy of two adjacent superpixels. The entropy can indicate the information content between two image patches. In other words, if the joint entropy  $h_{ij}$  is low, one image patch can be predicted by the other, which corresponds to low information [20]. The joint entropy is calculated as a sum of joint probability distribution of corresponding intensities of two superpixels:

$$h_{ij} = \sum_{a,b} P_{ij}(a, b) \log P_{ij}(a, b) \quad (4)$$

where  $P_{ij}(a, b)$  is the marginal probability distribution of two superpixels:

$$P_{ij}(a, b) = \frac{H_{ij}(a, b)}{\sum_{a=0}^{L-1} \sum_{b=0}^{M-1} H_{ij}(a, b)} \quad (5)$$

where  $H_{ij}(a, b)$  is the joint histogram of two image patches.  $L$  and  $M$  is the length and width of each image patch respectively. To calculate the joint entropy of two superpixels, the sizes of two superpixels need to be the same. Hence, we fit the compared superpixels in one minimum enclosing rectangle before calculation.

Specifically, we use the nodes on the image boundary as background seeds and construct four saliency maps using nodes on four sides of image as queries respectively.

As shown in Fig. 1, the final saliency map for stage one is integrated by multiplying four saliency maps:

$$S_c(i) = S_t(i) \times S_b(i) \times S_l(i) \times S_r(i) \quad (6)$$

where  $S_t(i)$ ,  $S_b(i)$ ,  $S_l(i)$  and  $S_r(i)$  is the saliency map using nodes on the top, bottom, left and right side of image as queries respectively. Taking the top as an example  $S_t(i)$  can be calculated by:

$$S_t(i) = 1 - \bar{f}^* \quad (7)$$

where  $\bar{f}^*$  denotes the normalized vector. The others can be calculated in the same way.

## 2.2 Saliency Optimization Using Depth Distribution

While most regions of salient objects are highlighted in the coarse saliency map computed from multi-feature-fused manifold ranking, some superpixels which belong to background are also highlighted. To solve this problem, we firstly segment depth map using an adaptive threshold to roughly select the superpixels of foreground. Since most regions of foreground have been highlighted in the coarse saliency map, the threshold is set to be the mean saliency of the coarse saliency map. The saliency less than mean saliency are set to be zero and otherwise is set to be one. Then we use this binary map  $S_{cb}(i)$  to segment depth map:

$$D_s = S_{cb} \times D \quad (8)$$

where  $D_s$  is the segmented depth map and  $D$  is the original depth map.

In physical world, salient objects usually have discrete depth values while background objects like walls, grounds or sky often have continuous depth values. In other words, the depth value of salient object often aggregate in a small interval. According to this finding, we use a nonlinear mapping function to refine depth value as a method of suppressing the depth value in background:

$$D'_s = 90 \arctan\left(\frac{1}{20}D_s - T_d\right) + \frac{255}{2} \quad (9)$$

where  $D'_s$  is the new depth map after mapping.  $T_d$  is the threshold value to control the enhancing interval and can be gained by choosing a reasonable distance for peak value of histogram.

At last, the saliency result of stage two  $S'_c$  is gained by:

$$S'_c = D'_s \times S_c \quad (10)$$

According to our observation, depth value does not naturally reflect the salient level of an object. An object with continuous depth like wall or ground usually has relatively close distance than salient object while their saliency is not that high. Thus, we use

Eq. (9) to guarantee that the utilization of depth value in stage two does not change the existing saliency of object. It is just used as a reference of segmentation to separate foreground objects with background.

### 2.3 Saliency-Depth Consistency Check

Saliency detection based on graph could be adversely affected if some background regions have relatively high contrast with its surrounding regions. Meanwhile, some parts of the saliency region could be assigned low saliency even if its surrounding parts have high saliency. These two weaknesses could result in inconsistency of salient regions, as Fig. 2 demonstrates. The saliency of some small regions in the background can be very high while the saliency of some object regions, especially those embedded in the object center can be low due to their low contrast to the four sides.

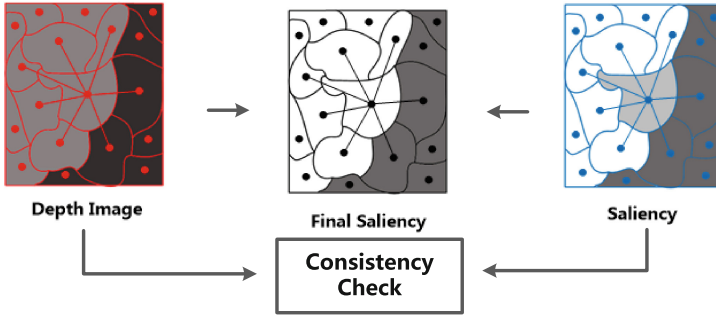


Fig. 2. Procedure of saliency-depth-consistency check

Since this problem cannot be solved by simply adjusting the weights of features in Eq. (3), a method of saliency-depth consistency check is introduced here to refine saliency map furtherly. The idea of this method is based on the fact that if regions share the same (or similar) depth value, their saliency should be the same (or similar) as well. According to this hypothesis, we firstly check if the saliency of a superpixel is consistent with its surrounding regions. If the saliency is inconsistent with its surrounding regions, we use a self-adaptive method to identify which of its surrounding superpixels are consistent with it according to their depth values. At last, we refine the inconsistent saliency with the mean of its consistent surrounding regions. We iterate all superpixels until all of them are consistent with its surrounding regions. The procedure of this algorithm is shown in Fig. 2.

In details, the saliency of superpixel  $S'_c(i)$  is defined to be consistent with its adjacent superpixels if:

$$\frac{1}{a} \frac{D_i}{D_{ij}} \leq \frac{S'_c(i)}{S'_c(j)} \leq a \frac{D_i}{D_{ij}}, \quad j \in N_i \quad (11)$$

where  $a > 1$  is the threshold factor.  $N_i$  is the set of superpixels which are adjacent to  $i$ .  $D_i$  and  $D_{ij}$  is the depth value of superpixel  $i$  and its adjacent superpixels respectively. In other words, if there exists a superpixel which does not meet Eq. (11), the saliency of superpixel  $i$  is inconsistent and should be refined by its adjacent superpixels. The new saliency of superpixel  $i$  can be calculated by:

$$S''_c(i) = \begin{cases} \text{Mean}(S'_c(j)) & \text{if } STD[\frac{D_i}{D_{i1}}, \frac{D_i}{D_{i2}} \dots \frac{D_i}{D_{ij}}] \leq T \\ \text{Mean}(S'_c(h)) & \text{if } STD[\frac{D_i}{D_{i1}}, \frac{D_i}{D_{i2}} \dots \frac{D_i}{D_{ij}}] > T \end{cases} \quad (12)$$

where  $\text{Mean}(S'_c(j))$  is the mean of all saliency of superpixels adjacent to  $i$ .  $\text{Mean}(S'_c(h))$  is the mean of saliency of superpixels whose depth are considered to be consistent with  $i$ . The threshold value  $T$  can be gained by minimizing  $\sigma$ :

$$\sigma = STD[\frac{D_i}{D_{i1}}, \frac{D_i}{D_{i2}}, \dots, \frac{D_i}{D_{ih}}] + STD[\frac{D_i}{D_{ih}}, \frac{D_i}{D_{i(h+1)}}, \dots, \frac{D_i}{D_{ij}}], j \in N_i$$

$$\frac{D_i}{D_{i1}} > \frac{D_i}{D_{i2}} > \dots, \frac{D_i}{D_{ih}} > \frac{D_i}{D_{i(h+1)}} > \dots, > \frac{D_i}{D_{ij}} \quad (13)$$

### 3 Experimental Results

To demonstrate the effectiveness of proposed method, we evaluate it on a RGB-D data set [18] captured by Microsoft Kinect which contains 1000 RGB-D images selected from 5000 natural images. Compared with other 3D data set, it has plenty of important characteristics such as high diversity and low bias. Moreover, both RGB and depth images in this data set share variable colour or depth ranges, which makes the salient object extremely difficult to be extracted by a simple thresholding.

We compared our algorithm with eight state-of-the-art saliency detection algorithms: ASD [16], DMC [18], SF [21], GS [22], MF [13], OP [23], PCA [24] and SWD [25]. Among these algorithms, both RGB and RGB-D, patch-based and graph-based methods were considered. The results of these algorithms are all provided by their original open source codes.

#### 3.1 Evaluation Measures

To conduct a quantitative performance evaluation, we select the method of precision-recall curve (PRC) and F-measure. The precision value indicates the proportion of salient pixels correctly assigned to all the pixels of detected regions, while the recall value correspond to the ratio of extract pixels be related to the ground-truth numbers. Like all state-of-the-art, the precision-recall curves are obtained by segmenting the saliency maps using threshold from 0 to 255. The F-measure is an overall performance measurement, which is defined as follow with the weighted harmonic of precision and recall:

$$F_{\beta} = \frac{(1 + \beta^2) \times \text{precision}}{\beta^2 \text{precision} + \text{Recall}} \quad (14)$$

where  $\beta^2 = 0.3$  is the emphasizing factor of precision.

To evaluate the precision of detected pixels and salient pixels in ground truth, the mean absolute error (MAE) [21] is introduced in this paper, which is defined as:

$$MAE = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H |s(x, y) - GT(x, y)| \quad (15)$$

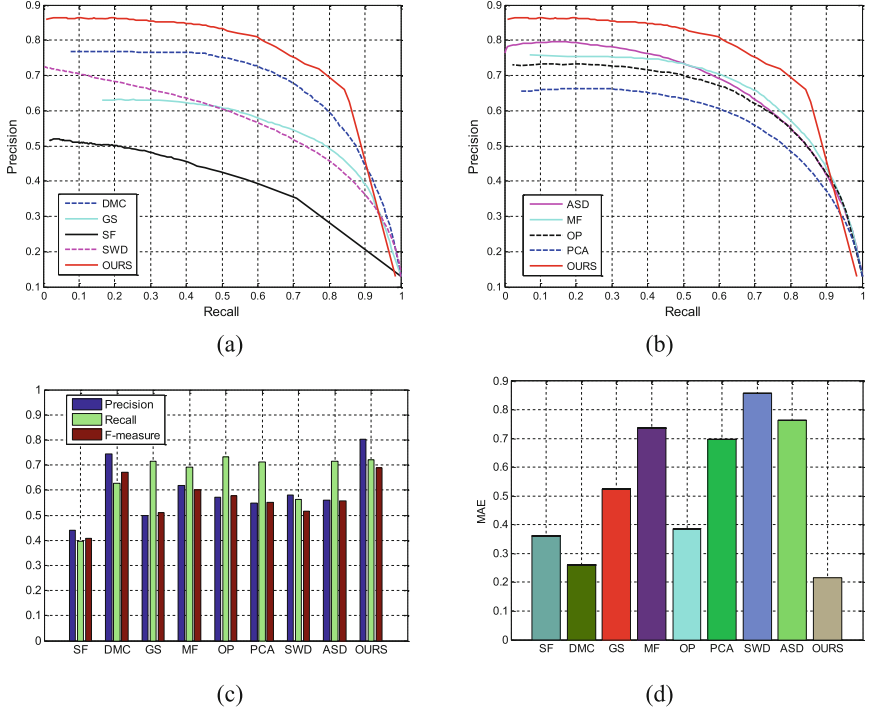
where  $W$  and  $H$  is the width and height of image.  $s(x, y)$  indicates the saliency map and  $GT(x, y)$  is the image of ground truth.

### 3.2 Quantitative Comparison with Other Methods

We examine our algorithm with eight state-of-the-art methods. The PR curve, F-measure and MAE of eight methods as well as ours are shown in Fig. 3. In saliency detection, the higher the precision and recall value is, the better performance of the method achieves. In this perspective, our algorithm has the best performance among all the eight methods despite some samples with high recall but relatively low precision. Although the precision-recall curve can give a qualitative evaluation of algorithms, the quantitative analysis still remains to be done. That is the reason we implement F-measure evaluation shown in Fig. 3(c). In that respect, the results of ours fully exceed the others except the recall is next only to OP. The MAE shown in Fig. 3(d) indicates SF, DMC and OP have relatively low MAE than other methods while our result is still the best among all eight methods. In detail, our final result (MAE = 0.216276) is around 17% better than the second performance DMC (MAE = 0.260925). This is mainly due to the saliency-depth consistency check which highlights the shadow part on object and removes some speckles in the background. In a word, as shown in Fig. 3 our algorithm is effective on the perspective of PR-curve on the RGB-D data base, only with the recall next to OP, but outperforms all the state-of-the-art methods in both F-measure and MAE.

It is worthwhile noticing that the performance of other methods measured by PR-curve, F-measure and MAE on this RGB-D data set are all poorer than the results presented in their original papers, which indicates that this RGB-D data set is more challenging compared with other data sets such as ASD and MSRA. Thus, throughout the evaluation on this RGB-D data set, it can be concluded that our algorithm can process more complex scenes and is more robust to the variation in the natural environments.

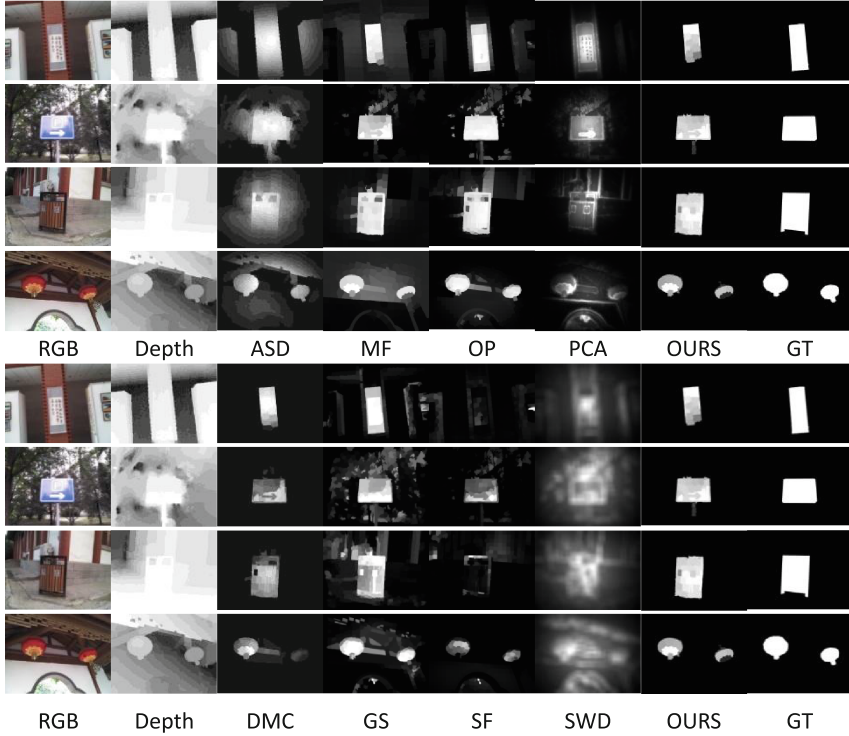




**Fig. 3.** Performance comparisons of proposed algorithm (a) PRC of proposed algorithm versus ASD [16], MF [13], OP [23] and PCA [24], (b) PRC of proposed algorithm versus DMC [18], GS [22], SF [21] and SWD [25], (c) F-measure scores, (d) MAE scores

### 3.3 Vision Comparison with Other Methods

The vision comparison of the results presented by our algorithm and eight state-of-the-art methods are shown in Fig. 4. It is obvious that 3D saliency detection methods like DMC can suppress the saliency of background more effectively. However, graph-based 2D saliency detection methods like MF and OP have stronger ability to highlight the saliency region and provide easier way to segment salient object from background. By taking advantage of both these two kinds of methods, our algorithm can highlight the intact boundary of object as well as suppress the complex interference in the background. Moreover, our algorithm can provide more inner-consistent saliency map due to the saliency-depth consistency check, which makes it much easier for further segmentation.



**Fig. 4.** Vision comparison of results computed by different methods. The proposed algorithm has better inner-consistency than others

## 4 Conclusion

In this paper, a novel three-stage saliency detection method using multi-feature-fused manifold ranking and depth optimization is presented. By fusing the depth information and joint entropy with color contrast in CIElab space, our algorithm is more robust in varying natural environments. To reduce the saliency of isolated background regions and highlight the saliency regions uniformly, a segment strategy as well as saliency-depth consistency check is proposed. Moreover, our algorithm takes advantage of both the graph-based saliency detection method and 3D saliency detection method, which makes detected salient objects have intact boundaries as well as inner-consistent saliency. Experimental results comparing with eight state-of-the-art approaches on a challenging RGB-D data set illustrates that the proposed method brings out desirable performance and produces more robust and visual favorable results in terms of the precision and recall analysis, F-measure and MAE. Despite of few limitations, our algorithm indeed promotes the quality of saliency detection in natural scenes and has a great application prospect.

**Acknowledgments.** The authors would like to thank the anonymous reviewers for their comments. This research was supported in part by the National Natural Science Foundation of China (Grant No. 61473144) and the Innovation Experiment Competition Funding for Graduates of Nanjing University of Aeronautics and Astronautics.

## References

1. 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)
2. Li, H., Wu, W., Wu, E.: Robust salient object detection and segmentation. In: Zhang, Y.-J. (ed.) *ICIG 2015. LNCS*, vol. 9219, pp. 271–284. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-21969-1\\_24](https://doi.org/10.1007/978-3-319-21969-1_24)
3. Schade, U., Meinecke, C.: Texture segmentation: do the processing units on the saliency map increase with eccentricity? *Vis. Res.* **51**(1), 1–12 (2011)
4. Fang, Y., et al.: Saliency detection in the compressed domain for adaptive image retargeting. *IEEE Trans. Image Process.* **21**(9), 3888–3901 (2012)
5. Chen, Y., Pan, Y., Song, M., et al.: Image retargeting with a 3D saliency model. *Signal Process.* **112**, 53–63 (2015)
6. Guo, C., Zhang, L.: A novel multi-resolution spatiotemporal saliency detection model and its applications in image and video compression. *IEEE Trans. Image Process.* **19**(1), 185–198 (2010)
7. Ren, Z., et al.: Region-based saliency detection and its application in object recognition. *IEEE Trans. Circuits Syst. Video Technol.* **24**(5), 769–779 (2014)
8. Zhu, J.Y., et al.: Unsupervised object class discovery via saliency-guided multiple class learning. In: *IEEE Conference on Computer Vision and Pattern Recognition 2012*, pp. 3218–3225. IEEE, Rhode Island (2012)
9. Sharma, G., Jurie, F., Schmid, C.: Discriminative spatial saliency for image classification. In: *IEEE Conference on Computer Vision and Pattern Recognition 2012*, pp. 3506–3513. IEEE, Rhode Island (2012)
10. Borji, A., et al.: Salient object detection: a survey. *Eprint Arxiv*, vol. 16, no. 7, pp. 1–26 (2014)
11. Cheng, M.M., et al.: Global contrast based salient region detection. In: *IEEE Conference on Computer Vision & Pattern Recognition 2015*, pp. 409–416. IEEE, Boston (2015)
12. Duan, L., et al.: Visual saliency detection by spatially weighted dissimilarity. In: *IEEE Conference on Computer Vision and Pattern Recognition 2011*, pp. 473–480. IEEE, Colorado (2011)
13. Yang, C., et al.: Saliency detection via graph-based manifold ranking. In: *IEEE Conference on Computer Vision and Pattern Recognition 2013*, pp. 3166–3173. IEEE, Oregon (2013)
14. Qin, Y., et al.: Saliency detection via cellular automata. In: *IEEE Conference on Computer Vision and Pattern Recognition 2015*, pp. 110–119. IEEE, Boston, (2015)
15. Lang, C., et al.: Depth matters: influence of depth cues on visual saliency. In: *European Conference on Computer Vision 2012*, pp. 101–115. IEEE, Florence (2012)
16. Ju, R., et al.: Depth saliency based on anisotropic center-surround difference. In: *IEEE International Conference on Image Processing 2014*, pp. 1115–1119. IEEE, Paris (2014)
17. Wang, A., et al.: Salient object detection with high-level prior based on Bayesian Fusion. *IET Comput. Vis.* **11**(2), 161–172 (2017)
18. Peng, H., et al.: RGBD salient object detection: a benchmark and algorithms. In: *European Conference on Computer Vision 2014*, pp. 92–109. IEEE, Zurich (2014)

19. Achanta, R., Shaji, A., Smith, K., et al.: SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(11), 2274–2282 (2012)
20. Hirschmuller, H.: Stereo processing by semiglobal matching and mutual information. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(2), 328–341 (2008)
21. Perazzi, F., et al.: Saliency filters: contrast based filtering for salient region detection. In: *IEEE Conference on Computer Vision and Pattern Recognition 2012*, pp. 733–740. IEEE, Rhode Island (2012)
22. Wei, Y., et al.: Geodesic saliency using background priors. In: *European Conference on Computer Vision 2012*, pp. 29–42. IEEE, Florence (2012)
23. Zhu, W., et al.: Saliency optimization from robust background detection. In: *IEEE Conference on Computer Vision and Pattern Recognition 2014*, pp. 2814–2821. IEEE, Columbus (2014)
24. Margolin, R., Tal, A., Zelnik-Manor, L.: What makes a patch distinct. *IEEE Trans. Comput. Vis. Pattern Recogn.* **9**(4), 1139–1146 (2013)
25. Duan, L., et al.: Visual saliency detection by spatially weighted dissimilarity. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 473–48. IEEE, Colorado, (2011)

Image and Graphics

9th International Conference, ICIG 2017, Shanghai,  
China, September 13-15, 2017, Revised Selected  
Papers, Part III

Zhao, Y.; Kong, X.; Taubmann, D. (Eds.)

2017, XVIII, 675 p. 344 illus., Softcover

ISBN: 978-3-319-71597-1