An Innovative Salient Object Detection Using Center-Dark Channel Prior

Chunbiao Zhu
SECE, Shenzhen Graduate School
Peking University
Shenzhen, China
zhuchunbiao@pku.edu.cn

Ge Li*
SECE, Shenzhen Graduate School
Peking University
Shenzhen, China
*geli@ece.pku.edu.cn

Wenmin Wang
SECE, Shenzhen Graduate School
Peking University
Shenzhen, China
wangwm@pkusz.edu.cn

Ronggang Wang
SECE, Shenzhen Graduate School
Peking University
Shenzhen, China
rgwang@pkusz.edu.cn

Abstract

Saliency detection aims to detect the most attractive objects in images, which has been widely used as a foundation for various multimedia applications. In this paper, we propose a novel salient object detection algorithm for RGB-D images using center-dark channel prior. First, we generate an initial saliency map based on a color saliency map and a depth saliency map of a given RGB-D image. Then, we generate a center-dark channel map based on a center saliency prior and a dark channel prior. Finally, we fuse the initial saliency map with the center dark channel map to generate the final saliency map. The proposed algorithm is evaluated on two public RGB-D datasets, and the experimental results show that our method outperforms the state-of-the-art methods.

1. Introduction

Saliency detection is a process of getting a visual attention region precisely from an image. The attention is a behavioral and cognitive process of selectively concentrating on one aspect within an environment while ignoring other things.

Early work on computing saliency aims to locate the visual attention region. Recently the field has been extended to locate and refine the salient regions and objects. Served as a foundation of various multimedia applications, salient object detection has been widely used in content-aware editing [3], image retrieval [4], object recognition [2, 13], object segmentation [8, 15], compression [10], image retargeting [22], etc.

In general, saliency detection algorithms mainly use top-down or bottom-up approaches. Top-down approaches are task-driven and need supervised learning. While bottom-up approaches usually use low-level cues without supervised learning, such as color features, distance features and other heuristic saliency features. One of the most used heuristic saliency feature [20, 1, 21, 14, 18, 23, 17, 6] is contrast, such as pixel-based or patch-based contrast, region-based contrast, multi-scale contrast, center-surround contrast, color spatial compactness, etc. Although those methods have their own advantages, they are not robust to specific situations which lead to inaccuracy results on challenging salient object detection datasets.

Recently, advances in 3D data acquisition techniques have motivated the adoption of structural features, improving the discrimination between different objects with the similar appearance. Although, depth cue can enhance salient object regions. It is very difficult to produce good results when a salient object has low depth contrast compared to its background.

Aiming to address the aforementioned difficulties of saliency detection on challenging datasets, in this paper, we propose an innovative saliency detection algorithm using center dark channel prior. We firstly generate an initial saliency map based on a color saliency map and a depth saliency map of a given RGB-D image. Second, since, salient objects are always located in the center of an image, and besides, we find that the dark channel prior can provide an additional cue for saliency detection. Therefore, we generate a novel center-dark channel map based on a center saliency prior and a dark channel prior to improve the performance. Finally, we fuse the initial saliency map with the center-dark channel prior to generate the final saliency map.

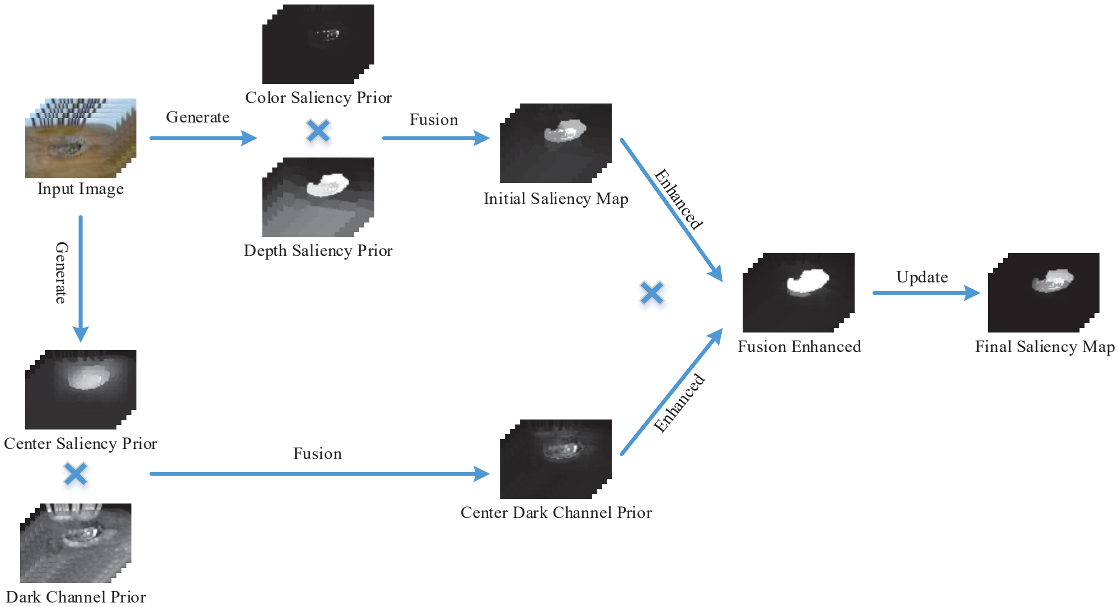


Figure 1. The framework of the proposed algorithm.

In summary, the main contributions of our work include:

1. We propose an innovative saliency detection algorithm to deal with the challenging scenarios.

2. We introduce a center dark channel prior for the first time in the saliency detection field to enhance the saliency detection performance.

3. We propose a direction for future research in the discussion part.

2. Related Work

RGB-D saliency computation is a rapidly growing field, offering object detection and attention prediction in a manner that is robust to appearance. Therefore, some algorithms [26, 24, 19, 5, 7, 25] adopt depth cues to deal with the challenging scenarios. In [26], Zhu et al. propose a framework based on cognitive neuroscience, and use depth cue to represent the depth of real scenario. In [5], Cheng et al. compute salient stimuli in both color and depth spaces. In [19], Peng et al. provide a simple fusion framework that combines existing RGB-based saliency with new depth-based saliency. In [7], Geng et al. define saliency using depth cue computed from stereo images. Their results show that stereo saliency is a useful consideration compared to previous visual saliency analysis. All of them demonstrate the effectivity of depth cue in improvement of salient object detection.

However, depth cue can not warrant robustness of saliency detection when a salient object has low depth contrast compared to the background. Inspired by cognitive neuroscience, we find that an image always possesses the salient objects in center position. As a result, many algorithms adopt center prior to enhance the salient regions. In [26], Zhu et al. propose a framework based on cognitive neuroscience, and use center prior to imitate the human central fovea. Furthermore, in [12], the contrast against image boundary is used as a new regional feature to enhance the center position. In [20], Qin et al. compare the boundary clusters with center clusters, and then generate different color distinction maps with complementary advantages and integrate them by taking spatial distance into the consideration. All of them demonstrate the center prior can strengthen the saliency regions.

On the other hand, dark channel prior, which was first put forward in [9], is used for single image haze removal. The dark channel prior is based on the statistics of outdoor haze-free images, which find that the dark pixels often have very low intensity in at least one color channel. Partially inspired by the well-known dark-object subtraction technique [11] widely used in multispectral remote sensing system. We figure out that salient objects tend to have the different transmissivity with the most regions of the background. As a result, we propose an algorithm to combine dark channel prior with saliency detection results to enhance saliency detection.

3. The Proposed Algorithm

3.1. Saliency map initialization

We initialize the saliency map by extracting color and depth features from the original image I_o and the depth map

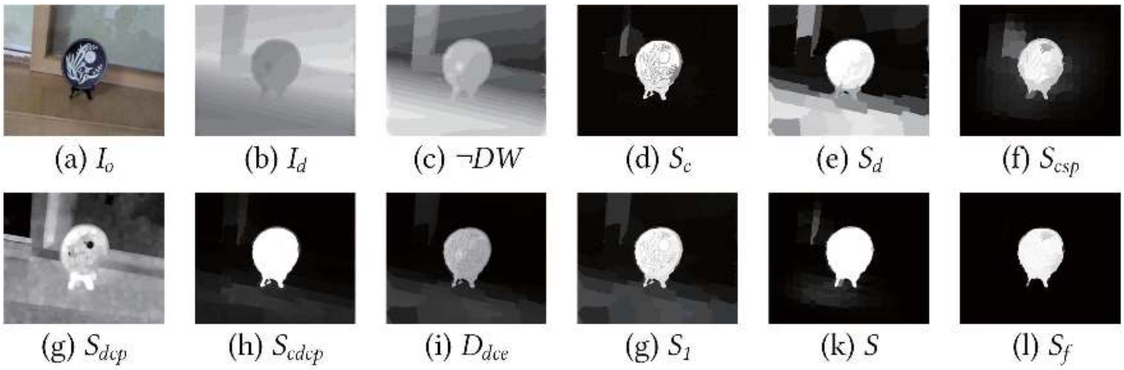


Figure 2. The visual process of the proposed algorithm.

I_d, respectively. First, the image I_o is segmented into K regions based on color via the K-means algorithm. Define:

$$S_c(r_k) = \sum_{i=1, i \neq k}^{K} P_i W_d(r_k) D_c(r_k, r_i), \tag{1}$$

where $S_c(r_k)$ is the color saliency of region k, $k \in [1, K]$, r_k and r_i represent regions k and i respectively, $D_c(r_k, r_i)$ is the Euclidean distance between region k and region i in L*a*b color space, P_i represents the area ratio of region r_i compared with the whole image, $W_d(r_k)$ is the spatial weighted term of the region k, set as:

$$W_d(r_k) = e^{\frac{-D_o(r_k, r_i)}{\sigma^2}}, \tag{2}$$

where $D_o(r_k, r_i)$ is the Euclidean distance between the centers of region k and i, σ is the parameter controlling the strength of $W_d(r_k)$.

Similar to color saliency, we define:

$$S_d(r_k) = \sum_{i=1, i \neq k}^{K} P_i W_d(r_k) D_d(r_k, r_i), \tag{3}$$

where $S_d(r_k)$ is the depth saliency of I_d, $D_d(r_k, r_i)$ is the Euclidean distance between region k and region i in depth space.

In most cases, a salient object is always located in the center of an image or close to a camera. Therefore, we assign the weights to both center-bias and depth for both color and depth images. The weight of the region k is set as:

$$W_{cd}(r_k) = \frac{G(\| P_k - P_o \|)}{N_k} DW(d_k), \tag{4}$$

where $G(\cdot)$ represents the Gaussian normalization, $\| \cdot \|$ is Euclidean distance, P_k is the position of the region k, P_o is the center position of this map, N_k is the number of pixels in region k, $DW(d_k)$ is the depth weight, which is set as:

$$DW(d_k) = (max\{d\} - d_k)^{\mu}, \tag{5}$$

where $max\{d\}$ represents the maximum depth of the image, and d_k is the depth value of region k, μ is a fixed value for a depth map, set as:

$$\mu = \frac{1}{max\{d\} - min\{d\}}, \tag{6}$$

where $min\{d\}$ represents the minimum depth of the image.

Then, the initial saliency value of the region k is calculated as:

$$S_1(r_k) = G(S_c(r_k) W_{cd}(r_k) + S_d(r_k) W_{cd}(r_k)). \tag{7}$$

3.2. Center-dark channel prior

We mix center saliency prior and dark channel prior to generate a new prior, we define it as the center dark channel prior. By using this prior, we can get a more accurate saliency map.

Center saliency prior. Inspired by cognitive neuroscience, human eyes use central fovea to locate objects and make them clearly visible. Therefore, most of the images taken by cameras always locate salient objects around the center. Aiming to get center saliency map, we use the B-SCA algorithm [20]. It constructs the global color distinction and the spatial distance matrix based on clustered boundary seeds and integrates them into a background-based map. Thus it can improve the accurate of center objects, erase the image edges' effect. As shown in the Fig. 2(f), the center saliency map can remove the surroundings of the image and reserve most of the salient regions. We denote this center-bias saliency map as S_{csp}.

Dark channel prior. The dark channel prior is a popular prior which is widely used in image haze removal field [9]. It is based on the statistics of outdoor haze-free images. The dark channel can detect the most haze-opaque region and improve the atmospheric light estimation. Inspired by dark channel prior, we find that the foreground and background have the different transmissivity, so, we can distinguish the salient objects from the backgrounds as shown in fig. 2(g). We apply this theory to saliency detection fields. And we denote the results map of the dark channel prior as S_{dcp}.

3.3. Saliency refinement with an updated fusion

Based on the initialized saliency map, we utilize an updated fusion to refine the saliency maps. The updated fusion consists of depth cue based enhancement, center-dark channel prior based enhancement and updated fusion based enhancement.

Depth cue based enhancement. Depth cue makes salient objects prominent. We denote:

$$D_{dce}(d_k) = Norm(\neg DW(d_k)), \tag{8}$$

where $\neg$ is the negation operation which can enhance the saliency degree of front regions as shown in Fig. 2(c), because the foreground object has low depth value in depth map while the background object has high depth value. $Norm(\cdot)$ is the normalized operation.

Centre-dark channel prior based enhancement. We combine the centre saliency prior and dark channel prior to enhance the final saliency results. Denote:

$$S_{cdcp}(r_k) = Norm(S_{csp}(r_k))Norm(S_{dcp}(r_k)), \tag{9}$$

Updated fusion based enhancement. We fuse depth cue based enhancement and centre-dark channel prior based enhancement with the initial saliency value. Denote:

$$S(r_k) = (1-e^{-(S_1(r_k)+D_{dce}(d_k)+S_{cdcp}(r_k))})S_1(r_k)S_{csp}(r_k), \tag{10}$$

where $S(r_k)$ is the fusion enhanced saliency value and $S_1(r_k)$ is the initial saliency value denoted in Section 2.1.

To refine the saliency results, we updated the fusion enhanced saliency value, denoted as:

$$S_f(r_k) = 1 - e^{-(S_1(r_k)S_{csp}(r_k)S(r_k))}. \tag{11}$$

where $S_f(r_k)$ is the final saliency value.

From the Fig. 2, we can see the visual results of the proposed algorithm. The main steps of the proposed salient object detection algorithm are summarized in Algorithm 1.

4. Experimental Evaluation

4.1. Datasets and evaluation indicators

Datasets. We evaluate the performance of the proposed saliency detection algorithm on two RGBD standard datasets: RGBD1* [5] and RGBD2* [19]. RGBD1* has 135 indoor images taken by Kinect with the resolution of 640×480. This dataset has complex backgrounds and irregular shapes of salient objects. RGBD2* contains 1000 images with two different resolutions of both 640×480 and 480×640, respectively.

Algorithm 1 The proposed saliency algorithm using centre-dark channel prior

Input: original maps I_o, depth maps I_d;
Output: final saliency values S_f;
1: **for** each region $k = 1, K$ **do:**
2: compute color saliency values $S_c(r_k)$ and depth saliency values $S_d(r_k)$;
3: calculate the center-bias and depth weights $W_{cd}(r_k)$;
4: get the initial saliency value $S_1(r_k)$;
5: **end for**
6: obtain centre saliency priors S_{csp} and dark channel priors S_{dcp};
7: figure out the final saliency values S_f by the updated fusion ;
8: **return** final saliency values S_f;

Evaluation indicators. Experimental evaluations are based on standard measurements including precision-recall curve, ROC curve and MAE (Mean Absolute Error). The MAE is formulated as:

$$MAE = \frac{\sum_{i=1}^{N} \|GT_i - S_i\|}{N}. \tag{12}$$

where N is the number of the testing images, GT_i is the area of the ground truth of image i, S_i is the area of detection result of image i.

4.2. Ablation study

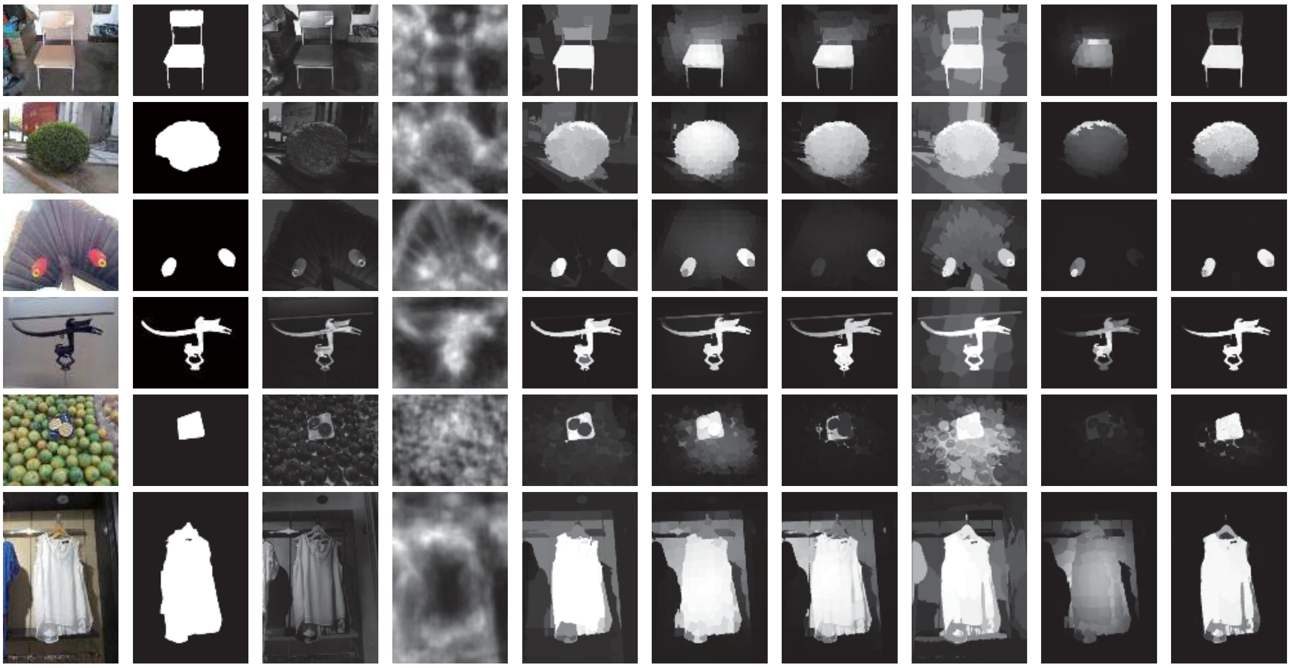
We first validate the effectiveness of each step in our method: initial saliency detection, depth cue enhanced saliency detection, centre-dark channel prior saliency detection, updated fusion saliency detection and final saliency detection. Table. 1 shows the validation results on RGB-D1* dataset. We can clear see the accumulated processing gains after each step, and the final saliency results shows a good performance. After all, it proves that each steps in our algorithm is effective for generating the final saliency maps.

Each Steps	S_1	D_{dce}	S_{cdcp}	S	S_f
MAE Values	0.2004	0.1685	0.1489	0.1343	0.0794

Table 1. MAE Values of each step in the proposed algorithm.

4.3. Comparison

To further illustrate the effectiveness of our algorithm, we compare our proposed methods with FT [1], SIM [18], HS [21], BSCA [20], LPS [14], RGBD1 [5], RGBD2 [19]. We use the codes provided by the author to reproduce their experiments. For all the compared methods, we use the default settings suggested by the authors. And for the Eq. 2,



(a) Image (b) GT (c) FT (d) SIM (e) HS (f) BSCA (g) LPS (h) RGBD1 (i) RGBD2 (j) OURS

Figure 4. Visual comparison of saliency maps on two datasets. (a) - (j) represent: original images, ground truth, FT, SIM, HS, BSCA, LPS, RGBD1, RGBD2 and OURS, respectively.

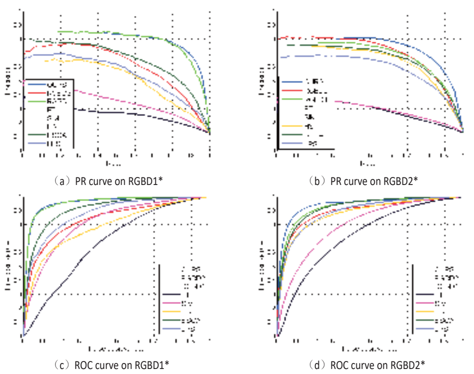


Figure 3. PR curve and ROC curve of different methods on two datasets.

Methods	RGBD1* Dataset	RGBD2* Dataset
FT	0.2049	0.2168
SIM	0.3740	0.3957
HS	0.1849	0.1909
BSCA	0.1851	0.1754
LPS	0.1406	0.1252
RGBD1	0.3079	0.3207
RGBD2	0.1165	0.1087
OURS	**0.0794**	**0.0860**

Table 2. MAE evaluation results.The best results are shown in boldface.

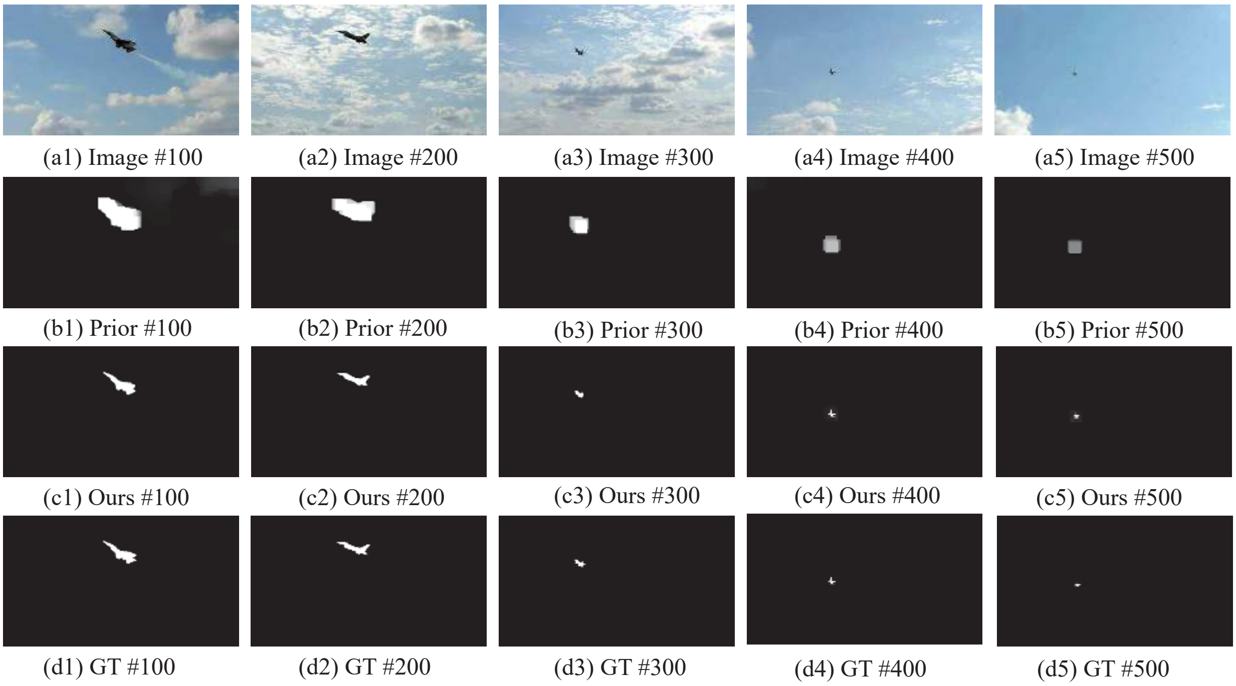
we take $\sigma^2 = 0.4$ which has the best contribution to the results.

The precision and recall evaluation results and ROC e-valuation results are shown in Fig. 3.

From the precision-recall curves and ROC curves, we can see that our saliency detection results can achieve better results on both RGBD1* and RGBD2* datasets.

MAE results on both RGBD1* and RGBD2* datasets are shown in Table 2, where the lower value, the better performance. The best results are shown in boldface. Compared with the MAE values, it can be observed that our saliency detection method is superior and can obtain the most precise salient regions than that of other approaches.

The visual comparisons are given in Fig. 4, which clearly demonstrate the advantages of our method. We can see that our method can detect both a single salient object and multiple salient objects more precisely. In contrast, the compared methods may fail in some situations.



(a1) Image #100 (a2) Image #200 (a3) Image #300 (a4) Image #400 (a5) Image #500

(b1) Prior #100 (b2) Prior #200 (b3) Prior #300 (b4) Prior #400 (b5) Prior #500

(c1) Ours #100 (c2) Ours #200 (c3) Ours #300 (c4) Ours #400 (c5) Ours #500

(d1) GT #100 (d2) GT #200 (d3) GT #300 (d4) GT #400 (d5) GT #500

Figure 5. The proposed algorithm is applied in small target detection. (a1)-(a5) represent different frames of original video.(b1)-(b5) represent different frames of the proposed priors detection results.(c1)-(c5) represent different frames of the proposed algorithm detection results. (d1)-(d5) represent different frames of the ground truth.

5. Discussion

It is very interesting to find that the proposed saliency detection algorithm using center-dark channel prior is also valid in small target detection by product. We present a part of our experimental results on the challenging dataset [16]. Small target detection plays an important role in many computer vision tasks, including early warning system, remote sensing and visual tracking. The experimental results by applying the proposed algorithm to small target detections are shown in Fig. 5, which support our claim. The behind reason why the proposed algorithm can be transplanted in small target detection is that (1) the proposed center-dark channel prior can locate the small objects and (2) the proposed saliency detection algorithm can refine the small targets. Therefore, we claim that the proposed center-dark channel prior is not only limited to saliency detection but also small target detection.

6. Conclusion

In this paper, we proposed an innovative saliency detection algorithm using center-dark channel prior. The proposed algorithm first detect the initial saliency maps based on color and depth cue. Then we figure out the center-dark channel saliency maps based on center saliency prior and dark channel prior. At last, we fuse them to get the final saliency maps by the updated fusion. The experiments' results show that the proposed algorithm outperforms the existing algorithms in both accuracy and robustness in different scenarios. Besides, by experiments, we claim that the proposed algorithm can also be applied to small object detection. To encourage future work, we make the source codes, experiment data, image datasets and other related materials public. All of these can be found on our project website[1].

7. Acknowledge

We would like to thank anonymous reviewers for their helpful comments on the paper. This work was supported by the grant of National Natural Science Foundation of China (No.U1611461), the grant of Science and Technology Planning Project of Guangdong Province, China (No.2014B090910001) and the grant of Shenzhen Peacock Plan (No.20130408-183003656).

[1]https://chunbiaozhu.github.io/ACVR2017/

References

[1] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-tuned salient region detection. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1597–1604, 2009.

[2] B. Alexe, T. Deselaers, and V. Ferrari. Measuring the objectness of image windows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2189–202, 2012.

[3] C. H. Chang, C. K. Liang, and Y. Y. Chuang. Content-aware display adaptation and interactive editing for stereoscopic images. *IEEE Transactions on Multimedia*, 13(4):589–601, 2011.

[4] M. M. Cheng, N. J. Mitra, X. Huang, and S. M. Hu. Salientshape: group saliency in image collections. *The Visual Computer*, 30(4):443–453, 2014.

[5] Y. Cheng, H. Fu, X. Wei, J. Xiao, and X. Cao. Depth enhanced saliency detection method. 55(1):23–27, 2014.

[6] A. Furnari, G. M. Farinella, and S. Battiato. An experimental analysis of saliency detection with respect to three saliency levels. In *Workshop at the European Conference on Computer Vision*, pages 806–821, 2014.

[7] Y. Geng. Leveraging stereopsis for saliency analysis. In *Computer Vision and Pattern Recognition*, pages 454–461, 2012.

[8] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 580–587, 2013.

[9] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1956–1963, 2009.

[10] L. Itti. *Automatic foveation for video compression using a neurobiological model of visual attention.* IEEE Press, 2004.

[11] P. S. C. Jr. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data . *Remote Sensing of Environment*, 24(3):459–479, 1988.

[12] T. Judd, K. Ehinger, F. Durand, and A. Torralba. Learning to predict where humans look. In *IEEE International Conference on Computer Vision, ICCV 2009, Kyoto, Japan, September 27 - October*, pages 2106–2113, 2009.

[13] M. Leo, G. Medioni, M. Trivedi, T. Kanade, and G. M. Farinella. Computer vision for assistive technologies. *Computer Vision and Image Understanding*, 154:1–15, 2017.

[14] H. Li, H. Lu, Z. Lin, and X. Shen. Inner and inter label propagation: Salient object detection in the wild. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 24(10):3176–3186, 2015.

[15] Y. Li, X. Hou, C. Koch, J. M. Rehg, and A. L. Yuille. The secrets of salient object segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 280–287, 2014.

[16] J. Lou, W. Zhu, H. Wang, and M. Ren. Small target detection combining regional stability and saliency in a color image. *Multimedia Tools and Applications*, pages 1–18, 2016.

[17] H. Lu, X. Zhang, J. Qi, T. Na, R. Xiang, and M. H. Yang. Cobootstrapping saliency. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, PP(99):1–1, 2016.

[18] N. Murray, M. Vanrell, X. Otazu, and C. A. Parraga. Saliency estimation using a non-parametric low-level vision model. 42(7):433–440, 2011.

[19] H. Peng, B. Li, W. Xiong, W. Hu, and R. Ji. *RGBD Salient Object Detection: A Benchmark and Algorithms*. Springer International Publishing, 2014.

[20] Y. Qin, H. Lu, Y. Xu, and H. Wang. Saliency detection via cellular automata. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 110–119, 2015.

[21] J. Shi, Q. Yan, X. Li, and J. Jia. Hierarchical image saliency detection on extended cssd. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(4):717, 2016.

[22] J. Sun and H. Ling. Scale and object aware image retargeting for thumbnail browsing. In *International Conference on Computer Vision*, pages 1511–1518, 2011.

[23] N. Tong, H. Lu, R. Xiang, and M. H. Yang. Salient object detection via bootstrap learning. In *Computer Vision and Pattern Recognition*, pages 1884–1892, 2015.

[24] C. Zhu, G. Li, X. Guo, W. Wang, and R. Wang. *A Multilayer Backpropagation Saliency Detection Algorithm Based on Depth Mining*, pages 14–23. Springer International Publishing, Cham, 2017.

[25] C. Zhu, G. Li, N. Li, X. Guo, W. Wang, and R. Wang. An innovative saliency detection framework with an example of image montage. In *ACM MultiMedia Workshop 2017 Submission Proposal for South African Academic Participation Proceedings*, 2017.

[26] C. Zhu, G. Li, W. Wang, and R. Wang. Salient object detection with complex scene based on cognitive neuroscience. In *Multimedia Big Data (BigMM), 2017 IEEE Third International Conference on*, pages 33–37. IEEE, 2017.