Local Background Enclosure for RGB-D Salient Object Detection

Anonymous CVPR submission

Paper ID 2328

Abstract

Recent work in salient object detection has considered the incorporation of depth cues from RGB-D images. In most cases, absolute depth, or depth contrast is used as the main feature. However, regions of high contrast in background regions cause false positives for such methods, as the background frequently contains regions that are highly variable in depth. Here, we propose a novel RGB-D saliency feature. Local background enclosure captures the spread of angular directions which are background with respect to the candidate region and the object that it is part of. We show that our feature improves over state-of-the-art RGB-D saliency approaches as well as RGB methods on the RGBD1000 and NJUS2000 datasets.

1. Introduction

Visual attention refers to the ability of the human visual system to rapidly identify scene components that stand out, or are salient, with respect to their surroundings. Early work on computing saliency aimed to model and predict human gaze on images [13]. Recently the field has expanded to include the detection of entire salient regions or objects [1][3]. These techniques have many computer vision applications, including compression [11], visual tracking [20], and image retargeting [19].

The saliency of a region is usually obtained by measuring contrast at a local [13] and/or global scale [8]. The majority of previous approaches compute contrast with respect to appearance-based features such as colour, texture, and intensity edges [14][7]. However, recent advances in 3D data acquisition techniques have motivated the adoption of structural features, improving discrimination between different objects with similar appearance.

RGB-D saliency methods typically incorporate depth directly, or use depth in a contrast measurement framework [26][12][27] [16][24], where contrast is computed as the difference between the means or distributions of foreground and background depth. Use of depth contrast in conjunction with colour contrast, various priors, and refinement schemes produces state-of-the-art results [26]. However, depth contrast is prone to false positives from background regions with large depth difference. Figure 1 illustrates some examples in which the foreground has relatively low contrast, making it challenging to detect using existing depth features. Contrast in background regions is unavoidable, and in general contrast in depth scenes can be dependent on random factors such as object placement and viewpoint. Although Ju *et al.* [16] has started to investigate the depth contrast for whole object structures, false positives still appear due to nearby regions with large depth difference as shown in Figure 1(f).

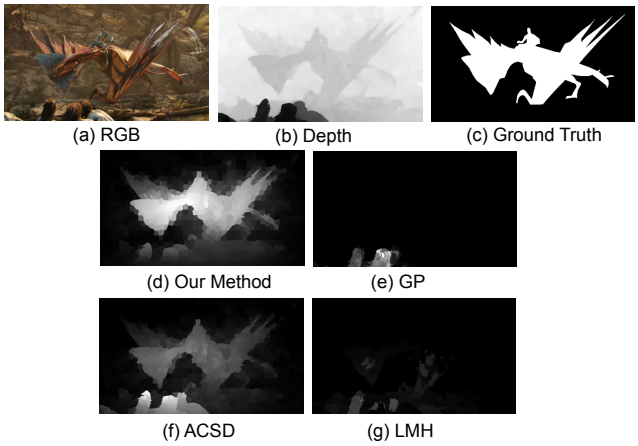


Figure 1. Saliency output on a depth image where foreground depth contrast is relatively low. Our method measures background enclosure of the object to overcome this problem.

Aiming to address this issue, we propose the Local Background Enclosure feature, which directly measures salient structure from depth. We note that salient objects tend to be characterised by being locally in front of surrounding regions, and the distance between an object and the background is not as important as the fact that the background surrounds the object for a large proportion of its boundary. The existence of background in a large spread of angular directions around the object implies pop-out structure and

thus high saliency. Conversely, background regions are less likely to exhibit pop-out structure. Thus we propose a depth saliency feature that incorporates two components. The first, which is proportional to saliency, is the angular density of background around a region, encoding the idea that a salient object is in front of most of its surroundings. The second feature component, which is inversely proportional to saliency, is the size of the largest angular region containing only foreground, since a large value implies significant foreground structure surrounding the object. This is the first time angular distributions of background directions have been explicitly incorporated for depth saliency. This feature is shown to be more robust than existing depth contrast-based measures. Further, we validate the proposed depth feature in a saliency system. We demonstrate that our depth feature out-performs state-of-the-art methods when combined with a depth prior, spatial prior, background prior, and Grabcut refinement.

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. Early works use depth as a prior to reweight 2D saliency maps [31][4][19]. These approaches do not consider relative depth, and work best when the range of salient objects is closer than the background.

More recently, the effectiveness of global contrast for RGB salient object detection [8] has inspired similar approaches for RGB-D saliency. Many existing methods measure global depth contrast, usually combined with colour and other modalities, to compute saliency [26][12][27][16][25][23]. While the majority of previous work computes depth contrast using absolute depth difference between regions, some methods instead use signed depth difference, improving results for salient objects in front of background [9]. Ju *et al*. [16] observe that while a salient object should be in front of its surrounds, patches on that object may be at a similar depth. However, as with other depth contrast methods, the primary feature of [16] is the depth difference between the foreground and background. Depth contrast methods are unlikely to produce good results when a salient object has low depth contrast compared to the rest of the scene (see Figure 1).

While depth contrast measurement forms the foundation of many approaches, it is common practice to enhance the resulting saliency maps by applying various priors and other refinement steps. The use of spatial and depth priors are widespread in existing work [25][16][6][12][27]. Ren *et al*. [26] explore orientation and background priors for detecting salient objects, and use PageRank and MRFs to optimize their saliency map. Peng *et al*. [25] incorporate object bias, and optimize their saliency map using a region growing approach. Ju et al [16] apply Grabcut segmentation to refine the boundaries of the generated saliency map.

Figure 2. Illustration of the local background sets (blue) for four different candidate regions (green). In this example the neighbourhood radius is $r = 200$ pixels, and the depth cutoff is $t = \sigma/2$. Note that patches lying on salient objects tend to be enclosed by the local background set.

3. Local Background Enclosure

In this section we introduce the Local Background Enclosure feature, denoted S, that quantifies the proportion of the object boundary that is in front of the background. S consists of an angular density component, F, and an angular gap component, G. The salient object detection system will be described in Section 4. Given an RGB-D image with pixel grid $I(x, y)$, we aim to segment the pixels into salient and non-salient pixels. For computational efficiency and to reduce noise from the depth image, instead of directly working on pixels, we oversegment the the image into a set of patches according to their RGB value. We denote the patches as $P \subset I$. We use SLIC [2] to obtain the superpixel segmentation, although our method is flexible to the type of segmentation method used.

3.1. Local Background Enclosure

Salient objects tend to be locally in front of their surroundings, and consequently will be mostly enclosed by a region of greater depth, as shown in Figure 2. We propose the Local Background Enclosure feature denoted by S based on depth. This is achieved by analysing angular density and gap statistics, denoted by F and G respectively, in order to quantify the proportion of the object boundary in front of the background.

3.1.1 Angular Density Component

We wish to measure the angular density of the regions surrounding P with greater depth than P, referred to as the local background. We consider a local neighbourhood N_P of P, consisting of all patches within radius r of P. That is, $N_P = \{Q \mid \|c_P - c_Q\|_2 < r\}$, where c_P and c_Q are patch centroids.

We define the local background $B(P, t)$ of P as the union of all patches within a neighbourhood N_P that have

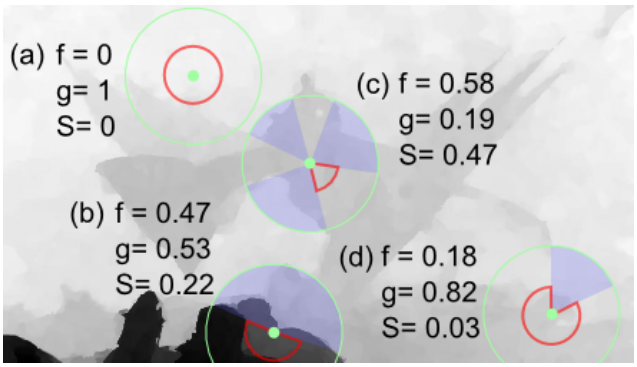


Figure 3. Illustration of the background enclosure feature computed at image locations marked by the green points with neighbourhood boundaries marked by a green line. The blue fill denotes angular regions containing points with greater depth than half a standard deviation from the center depth, with the maximum gap between these regions marked in red. The values of the angular density component f, the angular gap component g, and the final background enclosure saliency S are marked.

a mean depth above a threshold t from P.

$$B\left(P,t\right)=\bigcup\left\{P'\in N_P|D\left(P'\right)>D\left(P\right)+t\right\} \quad (1)$$

where $D\left(P\right)$ denotes the mean depth of pixels in P.

We define a function $f\left(P,B\left(P,t\right)\right)$ that computes the normalised ratio of the degree to which $B\left(P,t\right)$ encloses P.

$$f\left(P,B\left(P,t\right)\right)=\frac{1}{2\pi}\int_0^{2\pi}I\left(\theta,P,B\left(P,t\right)\right)\mathrm{d}\theta \quad (2)$$

where $I\left(\theta,P,B\left(P,t\right)\right))$ is an indicator function that equals 1 if the line passing through the centroid of patch P with angle θ intersects $B\left(P,t\right)$, and 0 otherwise. Note that we assume that P is compact. A visualisation of f is shown in Figure 3.

Thus $f\left(P,B\left(P,t\right)\right)$ computes the angular density of the background directions. Note that the threshold t for background is an undetermined function. In order to address this, as frequently used in probability theory, we employ the distribution function, denoted as F, instead of the density function f , to give a more robust measure. We define F as:

$$F(P)=\int_0^{\sigma}f\left(P,B\left(P,t\right)\right)\mathrm{d}t \quad (3)$$

where σ is the standard deviation of the mean patch depths within the local neighbourhood of P. This is given by $\sigma^2=\frac{1}{|B(P,t)|}\sum_{Q\in B(P,t)}\left(D(Q)-\overline{D}\right)^2$ where $\overline{D}=\frac{1}{|B(P,t)|}\sum_{Q\in B(P,t)}D(Q)$. This implicitly incorporates information about the distribution of depth differences between P and its local background.

3.1.2 Angular Gap Component

In addition to the angular density F, we introduce the angular gap statistic G. As shown in Figure 3, even though (c) and (b) have similar angular densities, we would expect (c) to have a significantly higher saliency since the background directions are more spread out. To capture this structure, we define the function $g\left(P,Q\right)$ to find the largest angular gap of Q around P and incorporate this into the saliency score.

$$g\left(P,Q\right)=\max_{(\theta_1,\theta_2)\in\Theta}\left\{|\theta_1-\theta_2|\right\} \quad (4)$$

where Θ denotes the set of boundaries (θ_1,θ_2) of angular regions that do not contain background:

$$\Theta=\left\{(\theta_1,\theta_2)\mid I\left(\theta,P,Q\right)=0\ \ \forall\theta\in[\theta_1,\theta_2]\right\} \quad (5)$$

A visualisation of g is shown in Figure 3.

Again, we use the distribution function of g, denoted by G:

$$G(P)=\int_0^{\sigma}1-\frac{g\left(P,B\left(P,t\right)\right)}{2\pi}\mathrm{d}t \quad (6)$$

The final Local Background Enclosure value is given by:

$$S(P)=F(P)\cdot G(P) \quad (7)$$

Figure 8 shows the generated saliency map on some example images. Note that the pop-out structure corresponding to salient objects is correctly identified. Depth contrast features fail to detect the objects, or exhibit high false positives.

4. Saliency Detection System

We construct a system for salient object detection using the proposed feature. Specifically, we reweight the Local Background Enclosure feature saliency using depth and spatial priors, and then refine the result using Grabcut segmentation. An overview of our system is given in Figure 4.

4.1. Depth, Spatial, and Background Prior

Studies report that absolute depth is an important component of pre-attentive visual attention, with closer objects more likely to appear salient to the human visual system [17]. Accordingly, scaling saliency by depth is a common refinement step in previous work [23], [6], [16], [28], [10], [5], [32][25][16][12][26]. We perform absolute depth reweighting using a depth prior $\mathcal{D}(x,y)$ to modulate the saliency of pixels with depth greater than the median depth of the image [16].

Another widely used prior is spatial bias, based on the tendency of the human visual system to fixate on objects near the center of an image [29]. A large number of existing saliency methods incorporate a center bias term to

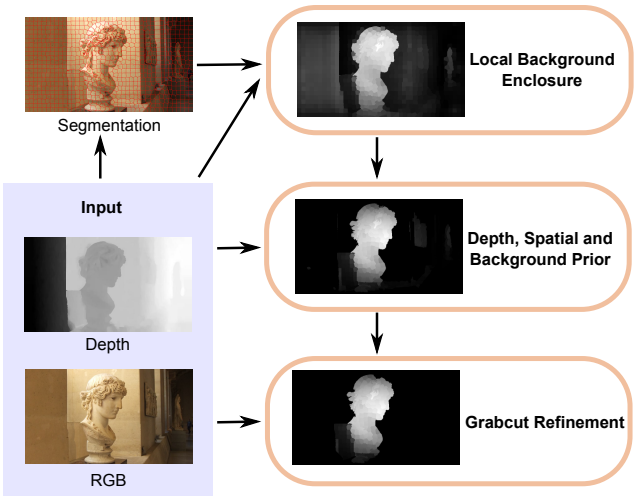


Figure 4. Overview of our saliency detection system. Given an RGB-D image and superpixel segmentation, we first compute our Local Background Enclosure feature, then apply depth, spatial, and center priors, and finally refine the result using Grabcut segmentation.

model this effect [25][16][6][12][27]. We incorporate this idea into our system, applying a Gaussian $\mathcal{G}(x, y)$ to reweight patch saliency based on the distance between the pixel (x, y) and the image center.

Recent works also incorporate a background prior based on some measure of boundary connectedness to improve detector precision [24][26]. We use the background prior map $\mathcal{B}(x, y)$ described in [33] to reweight saliency.

The low-level saliency map with priors applied is thus given by:

$$S_b = S \cdot \mathcal{D} \cdot \mathcal{G} \cdot \mathcal{B} \tag{8}$$

4.2. Grabcut Segmentation

The saliency map S_b may contain inaccurate foreground boundaries for parts of the object that do not exhibit strong pop-out structure. Boundary refinement is a common post-processing step employed in existing salient object detection systems (e.g. [6][25][21][26][12]). Similar to [22], we use Grabcut based boundary refinement to improve object boundaries using appearance information. The foreground model is initialized with a binary mask obtained by applying a threshold α_0 to S_b. The output Grabcut segmentation mask A is used to prune non-foreground areas from S_b. The refined saliency map is thus given by

$$S_g = A \cdot S_b \tag{9}$$

4.3. Implementation Details

The discrete version of the angular density function f is implemented using a histogram-based approximation, denoted as $\tilde{f}$. Let $h\left(i, P, B\left(P, t\right)\right)$ be an n bin polar occupancy histogram, where bin i is 1 if the corresponding angular range contains an angle between the centroids of P and a patch in $B\left(P, t\right)$, and 0 otherwise. We set $\tilde{f}$ to be equal to the fill ratio of h.

$$\tilde{f} = \frac{1}{n} \sum_{i=1}^{n} h\left(i, P, B\left(P, t\right)\right) \tag{10}$$

The distribution function F is computed numerically using $\tilde{F}$ by sampling $\tilde{f}$ at m equally spaced points across the integration range such that:

$$F(P) = \frac{1}{m} \sum_{i=1}^{m} \tilde{f}\left(P, B\left(P, \frac{i \cdot \sigma}{m}\right)\right). \tag{11}$$

Similarly, we define $\tilde{G}$ to evaluate G:

$$\tilde{G}(P) = \frac{1}{m} \sum_{i=1}^{m} 1 - \frac{1}{2\pi} \cdot g\left(P, \frac{i \cdot \sigma}{m}\right). \tag{12}$$

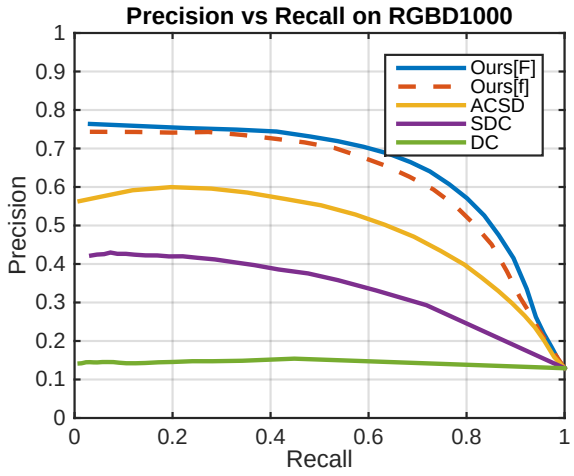
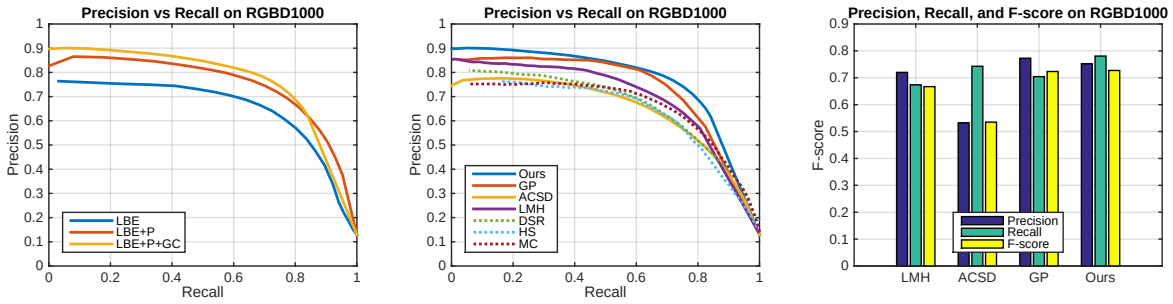


Figure 5. Performance of our depth feature (Ours[F]) plotted against performance of contrast-based depth features: depth contrast (DC), signed depth contrast (SDC), and ACSD [16] on RGBD1000. Also plotted is our feature using the density functions f and g with $t = \sigma/2$ (Ours[f]).

5. Experiments

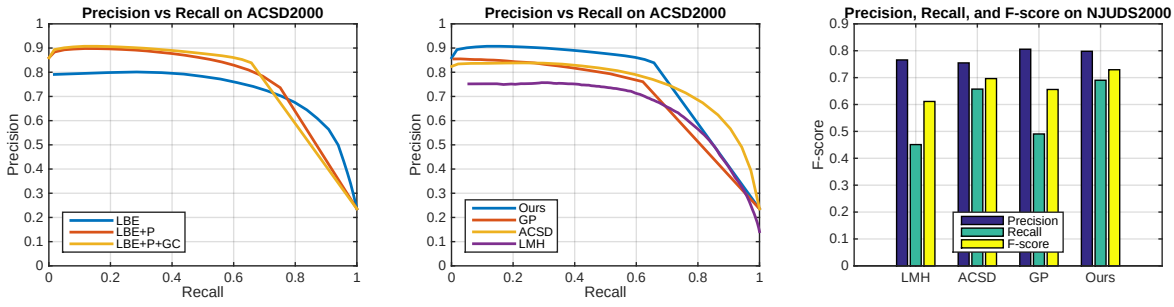
The performance of our saliency system is evaluated on two datasets for RGB-D salient object detection. RGBD1000 [25] contains 1000 structured light depth and RGB images. NJUDS2000 [16] contains 2000 disparity and RGB images computed from stereo image pairs.

The proposed Local Background Enclosure feature is compared against three types of depth-contrast based features on RGBD1000. These are: depth-contrast (DC) computed using KDE [25]; depth contrast with signed depth



(a) (b) (c)

Figure 6. Quantitative comparisons of performance over RGBD1000 dataset. (a) PR curves showing the effect of each component of the saliency system. P and GC refer to prior application and Grabcut refinement respectively. (b) PR curve of our saliency system against state-of-the-art RGB-D saliency systems. (c) F-measure of the different systems.



(a) (b) (c)

Figure 7. Quantitative comparisons of performance over NJUDS2000 dataset. (a) PR curves showing the effect of each component of the saliency system. P and GC refer to prior application and Grabcut refinement respectively. (b) PR curve of our saliency system against state-of-the-art RGB-D saliency systems. (c) F-measure of the different systems.

(SDC) such that patches with lower average depth do not contribute to the contrast score; and ACSD [13]. Additionally, in order to verify the contribution of using the distribution functions, we compute the product of the density functions $f(P,t) \cdot g(P,t)$ with fixed threshold $t = \sigma/2$.

We then evaluate the contribution of prior application and Grabcut refinement on our salient object detection system on both datasets. Finally, we compare our salient object detection system with three state-of-the-art RGB-D salient object detection systems: LMH [25], ACSD [16], and a recently proposed method that exploits global priors, which we refer to as GP [26]. We also include comparisons with three state-of-the-art 2D saliency algorithms DSR [18], HS [30], and MC [15] on RGBD1000.

5.1. Evaluation Metrics

We present the precision-recall curve and mean F-score to evaluate algorithm performance. The F-score is computed from the saliency output using an adaptive threshold equal to twice the mean of the image [1]. Note that the F-score is calculated as:

$$F_\beta = \frac{(1+\beta^2) \times Precision \times Recall}{\beta^2 \times Precision + Recall} \quad (13)$$

where $\beta = 0.3$ to weigh precision more than recall [1].

5.2. Experimental Setup

We set $n = 32$ histogram bins and $m = 10$ evaluation steps in our implementation of F and G respectively. These two values were found to provide a good trade-off between accuracy and efficiency for general use. The radius of the neighbourhood N_P should be set to equal the expected radius of the largest object to detect, thus we set it to half the image diagonal for general use. We use SLIC [2] on the colour image to generate the set of patches, with the number of patches set to the length of the diagonal of the image in pixels.

Our saliency method has one parameter - the threshold α_0 used to generate the foreground mask for Grabcut initialisation. We empirically set this to $\alpha_0 = 0.8$ in the experiments.

5.3. Results

Figure 5 shows that our method performs significantly better than the depth contrast based methods. It also demonstrates that using the distribution function gives improved results compared to using the density functions evaluated at fixed threshold t.

Our saliency system gives the highest F-score on both datasets, with GP providing the second best performance. We also produce the overall best PR curve on both datasets. This demonstrates that our feature is able to identify salient structure from depth more effectively than existing contrast-based methods.

Figure 8 shows the output of our salient detection system compared with state-of-the-art methods. Note that the other methods tend to have a high number of false positives due to depth contrast in background regions, for example depth change across a flat table is registered as salient by ACSD in the second row. The angular statistics employed by our depth feature provide a more robust measure of salient structure.

6. Conclusion

In this paper, we have proposed a novel depth feature that exploits depth background enclosure to detect salient objects in RGB-D images. We incorporate this feature into a salient object detection system using depth prior, spatial prior, and Grabcut refinement. Our approach out-performs existing methods on two publicly available RGB-D salient object detection datasets.

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. IEEE, 2009. 1, 5

[2] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(11):2274–2282, 2012. 2, 5

[3] A. Borji, D. N. Sihite, and L. Itti. Salient object detection: A benchmark. In *Computer VisionECCV 2012*, pages 414–429. Springer, 2012. 143. 1

[4] C. Chamaret, S. Godeffroy, P. Lopez, and O. Le Meur. Adaptive 3d rendering based on region-of-interest. In *SD&A*, 2010. 2

[5] C. Chamaret, S. Godeffroy, P. Lopez, and O. Le Meur. Adaptive 3d rendering based on region-of-interest. In *IS&T/SPIE Electronic Imaging*, pages 75240V–75240V. International Society for Optics and Photonics, 2010. 20. 3

[6] F. Chen, C. Lang, S. Feng, and Z. Song. Depth Information Fused Salient Object Detection. In *Proceedings of International Conference on Internet Multimedia Computing and Service*, ICIMCS '14, pages 66:66–66:70, New York, NY, USA, 2014. ACM. 2, 3, 4

[7] M.-M. Cheng, J. Warrell, W.-Y. Lin, S. Zheng, V. Vineet, and N. Crook. Efficient salient region detection with soft image abstraction. In *Computer Vision (ICCV), 2013 IEEE International Conference on*, pages 1529–1536. IEEE, 2013. 44. 1

[8] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu. Global contrast based salient region detection. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 409–416. IEEE, 2011. 1, 2

[9] Y. Cheng, H. Fu, X. Wei, J. Xiao, and X. Cao. Depth Enhanced Saliency Detection Method. In *Proceedings of International Conference on Internet Multimedia Computing and Service*, ICIMCS '14, pages 23:23–23:27, New York, NY, USA, 2014. ACM. 0. 2

[10] K. Desingh, K. M. Krishna, D. Rajan, and C. Jawahar. Depth really matters: Improving visual salient region detection with depth. In *Proc. BMVC*, 2013. 6. 3

[11] C. Guo and L. Zhang. A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression. *Image Processing, IEEE Transactions on*, 19(1):185–198, 2010. 1

[12] J. Guo, T. Ren, J. Bei, and Y. Zhu. Salient Object Detection in RGB-D Image Based on Saliency Fusion and Propagation. In *Proceedings of the 7th International Conference on Internet Multimedia Computing and Service*, ICIMCS '15, pages 59:1–59:5, New York, NY, USA, 2015. ACM. 1, 2, 3, 4

[13] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on pattern analysis and machine intelligence*, 20(11):1254–1259, 1998. a lot. 1

[14] B. Jiang, L. Zhang, H. Lu, C. Yang, and M.-H. Yang. Saliency detection via absorbing markov chain. In *Computer Vision (ICCV), 2013 IEEE International Conference on*, pages 1665–1672. IEEE, 2013. 27. 1

[15] B. Jiang, L. Zhang, H. Lu, C. Yang, and M.-H. Yang. Saliency detection via absorbing markov chain. In *ICCV*, 2013. 5

[16] R. Ju, L. Ge, W. Geng, T. Ren, and G. Wu. Depth saliency based on anisotropic center-surround difference. *ICIP. IEEE*, 2014. 1. 1, 2, 3, 4, 5

[17] C. Lang, T. V. Nguyen, H. Katti, K. Yadati, M. Kankanhalli, and S. Yan. Depth matters: Influence of depth cues on visual saliency. In *Computer VisionECCV 2012*, pages 101–115. Springer, 2012. 37. 3

[18] X. Li, H. Lu, L. Zhang, X. Ruan, and M.-H. Yang. Saliency detection via dense and sparse reconstruction. In *ICCV*, 2013. 5

[19] W.-Y. Lin, P.-C. Wu, and B.-R. Chen. Image Retargeting Using Depth Enhanced Saliency. *Proceedings of 3DSA2013*, 7:1, 2013. 2. 1, 2

[20] V. Mahadevan and N. Vasconcelos. Saliency-based discriminant tracking. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1007–1013. IEEE, 2009. 1

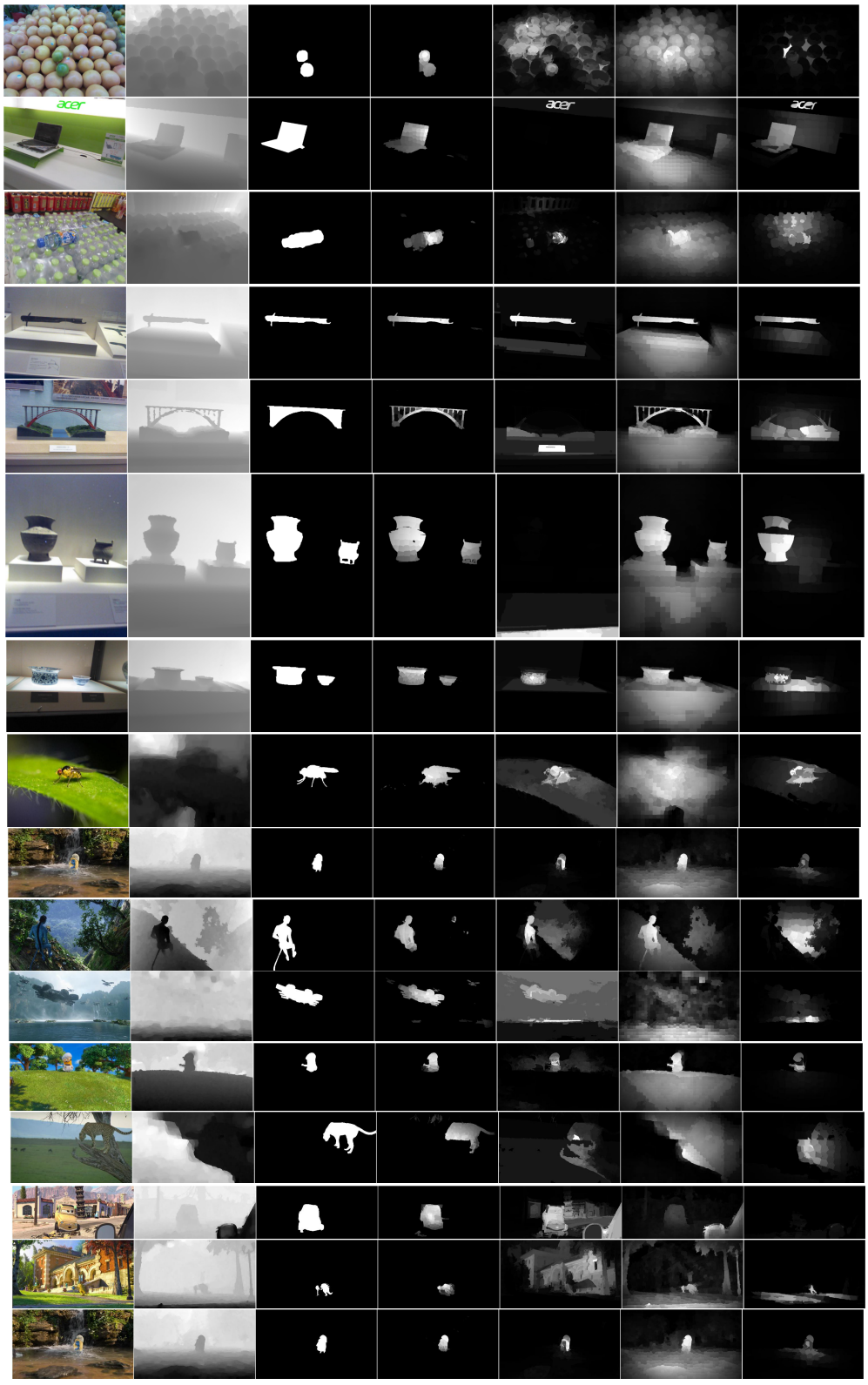


Figure 8. Comparison of output saliency maps. Left to right: rgb, depth, groundtruth, ours, GP, ACSD, LMH

[21] P. Mehrani and O. Veksler. Saliency Segmentation based on Learning and Graph Cut Refinement. In *BMVC*, pages 1–12. Citeseer, 2010. 20. 4

[22] P. Mehrani and O. Veksler. Saliency Segmentation based on Learning and Graph Cut Refinement. In *BMVC*, 2010. 4

[23] Y. Niu, Y. Geng, X. Li, and F. Liu. Leveraging stereopsis for saliency analysis. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 454–461. IEEE, 2012. 30. 2, 3

[24] H. Peng, B. Li, W. Xiong, W. Hu, and R. Ji. Rgbd salient object detection: a benchmark and algorithms. In *Computer Vision–ECCV 2014*, pages 92–109. Springer, 2014. 1, 4

[25] H. Peng, B. Li, W. Xiong, W. Hu, and R. Ji. Rgbd salient object detection: a benchmark and algorithms. In *Computer VisionECCV 2014*, pages 92–109. Springer, 2014. 3. 2, 3, 4, 5

[26] J. Ren, X. Gong, L. Yu, W. Zhou, and M. Yang. Exploiting Global Priors for RGB-D Saliency Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 25–32, 2015. 1, 2, 3, 4, 5

[27] H. Song, Z. Liu, H. Du, G. Sun, and C. Bai. Saliency detection for rgbd images. In *Proceedings of the 7th International Conference on Internet Multimedia Computing and Service*, page 72. ACM, 2015. 1, 2, 4

[28] Y. Tang, R. Tong, M. Tang, and Y. Zhang. Depth incorporating with color improves salient object detection. *The Visual Computer*, pages 1–11, Jan. 2015. 0. 3

[29] P.-H. Tseng, R. Carmi, I. G. Cameron, D. P. Munoz, and L. Itti. Quantifying center bias of observers in free viewing of dynamic natural scenes. *Journal of vision*, 9(7):4, 2009. 3

[30] Q. Yan, L. Xu, J. Shi, and J. Jia. Hierarchical saliency detection. In *CVPR*, 2013. 5

[31] Y. Zhang, G. Jiang, M. Yu, and K. Chen. Stereoscopic Visual Attention Model for 3d Video. In *Advances in Multimedia Modeling*, Lecture Notes in Computer Science, pages 314–324. Springer Berlin Heidelberg, 2010. 2

[32] Y. Zhang, G. Jiang, M. Yu, and K. Chen. Stereoscopic Visual Attention Model for 3d Video. In S. Boll, Q. Tian, L. Zhang, Z. Zhang, and Y.-P. P. Chen, editors, *Advances in Multimedia Modeling*, number 5916 in Lecture Notes in Computer Science, pages 314–324. Springer Berlin Heidelberg, 2010. 51. 3

[33] W. Zhu, S. Liang, Y. Wei, and J. Sun. Saliency optimization from robust background detection. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 2814–2821. IEEE, 2014. 4