Visual Saliency Based on Multiscale Deep Features
Supplementary Material

Guanbin Li Yizhou Yu
Department of Computer Science, The University of Hong Kong
https://sites.google.com/site/ligb86/mdfsaliency/

1. Quantitative comparison of different methods on SED dataset

SED dataset contains two subsets: SED1 and SED2. SED1 has 100 images each containing only one salient object while SED2 has 100 images each containing two salient objects. As shown in the first row of Fig. 1, our method achieves the highest precision in almost the entire recall range on both the SED1 and SED2 datasets.

Precision, recall and F-measure results using the adaptive threshold are shown in the second row of Figure 1, sorted by the F-measure. Our method also achieves the best performance on the overall F-measure as well as significant increases in both precision and recall. On SED1 dataset, our method achieves 87.67% precision and 78.95% recall while the second best (MR[8]) achieves 87.23% precision and 64.51% recall. On SED2 dataset, our method achieves 81.28% precision and 79.78% recall while the second best (wCtr*[9]) achieves 81.18% precision and only 71.54% recall. We can see that though the precision of other methods including MR[8], DRFI[4] and wCtr*[9] is comparable to ours, their recalls are much lower.

The third row of Fig. 1 shows the performance of our method compares to others in terms of the MAE measure. Our method successfully lowers the MAE by 14.15% with respect to the second best algorithm (wCtr*) on the SED1 dataset. On SED2 dataset, our method lowers the MAE by 8.76% with respect to the second best algorithm.

2. Part of the challenging images selected from our dataset

Fig. 2 and Fig. 3 show sample images from our own dataset (HKU-IS).

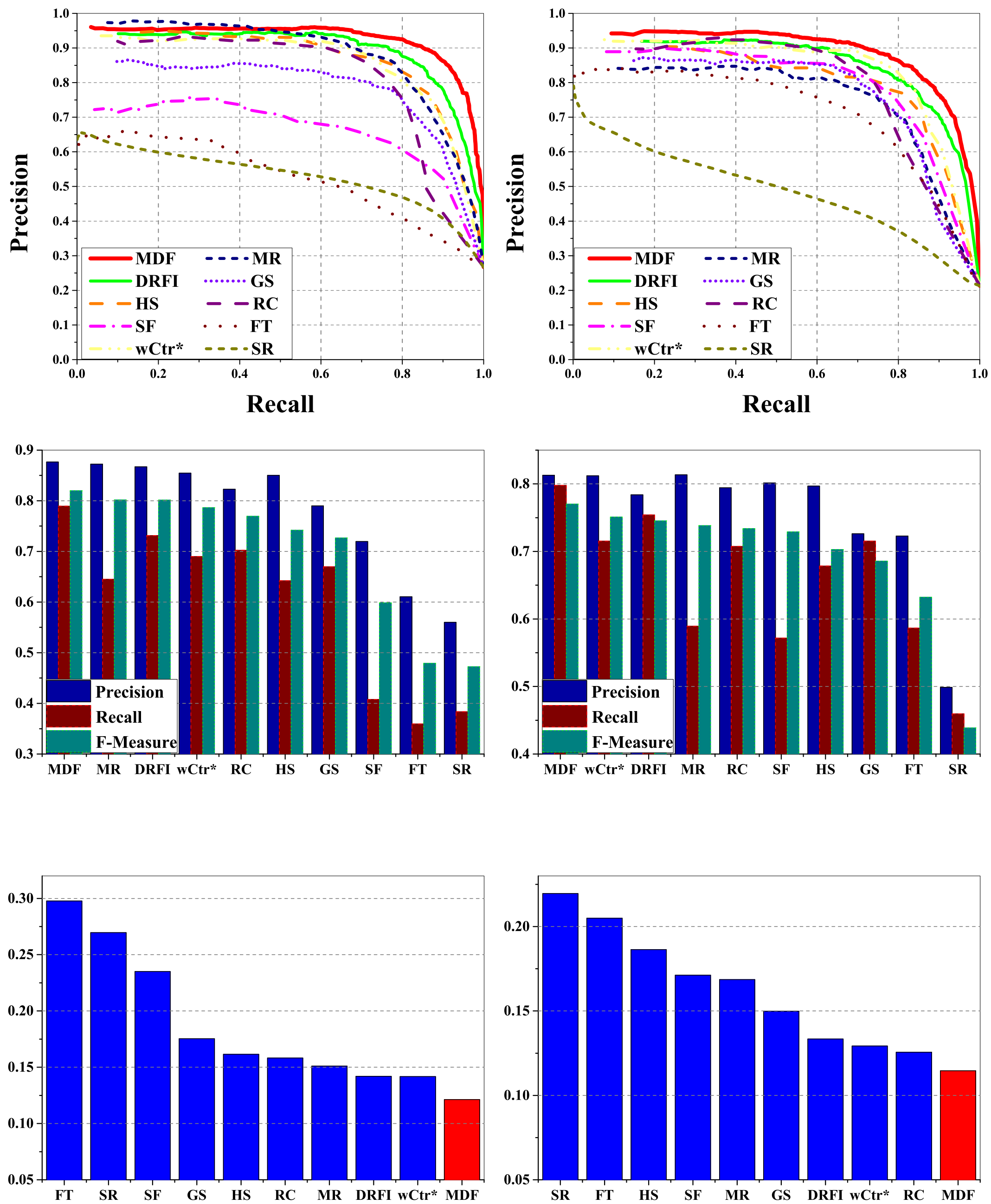
3. Visual comparison of saliency maps generated from 10 different methods

Fig. 4 and Fig. 5 demonstrate more visual comparison of saliency maps generated from 10 different methods, including ours (MDF). The orginal images are all selected from our dataset. The ground truth (GT) is shown in the last column. MDF consistently produces saliency maps closest to the ground truth. We compare MDF against spectral residual (SR[3]), frequency-tuned saliency (FT [1]), saliency filters (SF [5]), geodesic saliency (GS [6]), hierarchical saliency (HS [7]), regional based contrast (RC [2]), manifold ranking (MR [8]), optimized weighted contrast (wCtr* [9]) and discriminative regional feature integration (DRFI [4]).

References

[1] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-tuned salient region detection. In *CVPR*, 2009. 1, 5, 6
[2] M.-M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr, and S.-M. Hu. Global contrast based salient region detection. *TPAMI*, 2014. 1, 5, 6
[3] X. Hou and L. Zhang. Saliency detection: A spectral residual approach. In *CVPR*, 2007. 1, 5, 6
[4] H. Jiang, J. Wang, Z. Yuan, Y. Wu, N. Zheng, and S. Li. Salient object detection: A discriminative regional feature integration approach. In *CVPR*, 2013. 1, 5, 6
[5] F. Perazzi, P. Krahenbuhl, Y. Pritch, and A. Hornung. Saliency filters: Contrast based filtering for salient region detection. In *CVPR*, 2012. 1, 5, 6
[6] Y. Wei, F. Wen, W. Zhu, and J. Sun. Geodesic saliency using background priors. In *ECCV*. 2012. 1, 5, 6
[7] Q. Yan, L. Xu, J. Shi, and J. Jia. Hierarchical saliency detection. In *CVPR*, 2013. 1, 5, 6
[8] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang. Saliency detection via graph-based manifold ranking. In *CVPR*, 2013. 1, 5, 6
[9] W. Zhu, S. Liang, Y. Wei, and J. Sun. Saliency optimization from robust background detection. In *CVPR*, 2014. 1, 5, 6

978-1-4673-6964-0/15/$31.00 ©2015 IEEE



(a) (b)

Figure 1: Quantitative comparison of saliency maps generated by different methods on SED dataset. From left to right:(a) the SED1 dataset (b) the SED2 dataset. From top to bottom: the precision-recall curves of different methods, the precision, recall and F-measure using an adaptive threshold, and the mean absolute error.

Figure 2: Sample images in our dataset.

Figure 3: Sample images in our dataset.

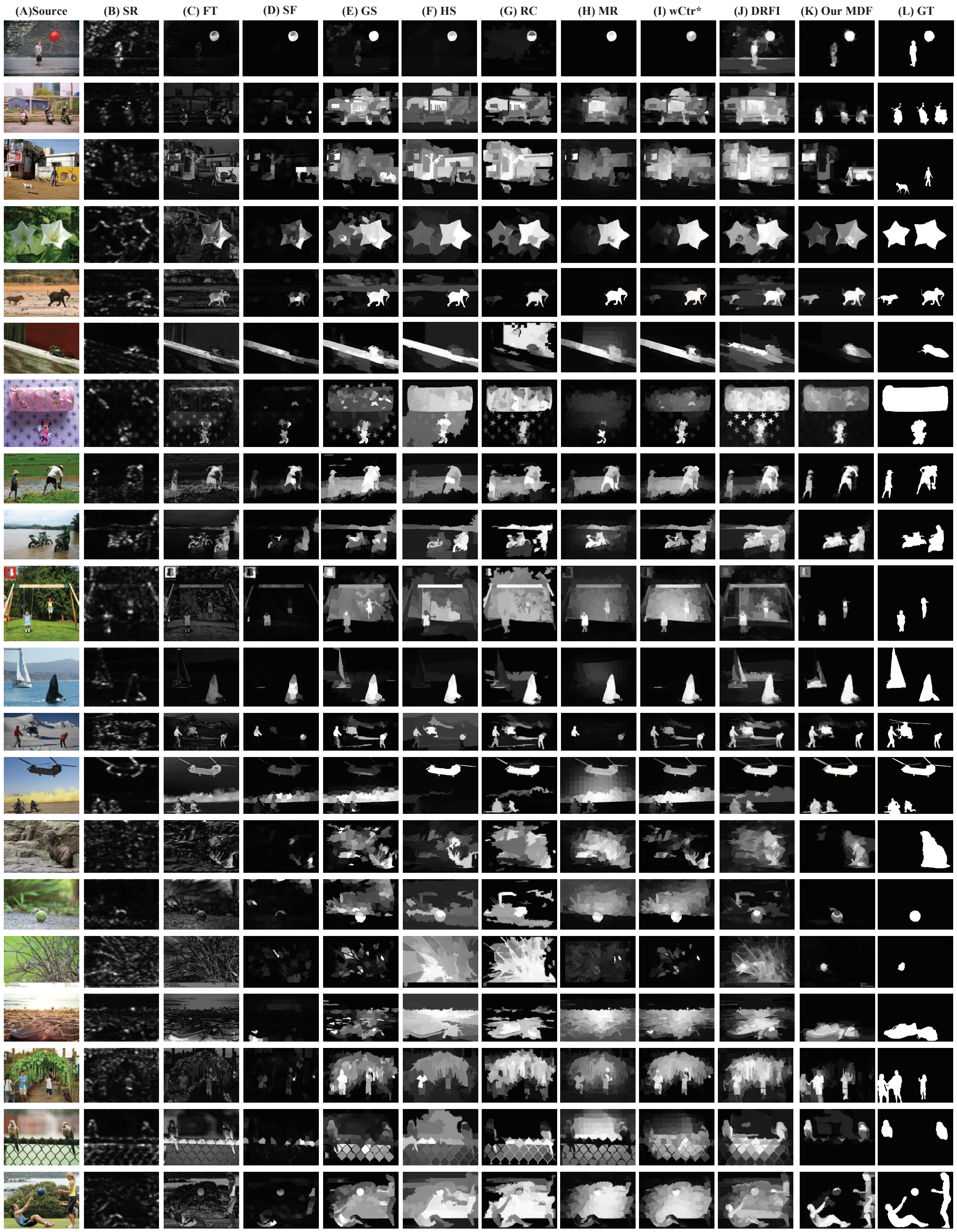


Figure 4: Visual comparison of saliency maps generated from 10 different methods, including ours (MDF). The ground truth (GT) is shown in the last column. We compare MDF against spectral residual (SR[3]), frequency-tuned saliency (FT [1]), saliency filters (SF [5]), geodesic saliency (GS [6]), hierarchical saliency (HS [7]), regional based contrast (RC [2]), manifold ranking (MR [8]), optimized weighted contrast (wCtr* [9]) and discriminative regional feature integration (DRFI [4]).

Figure 5: Visual comparison of saliency maps generated from 10 different methods, including ours (MDF). The ground truth (GT) is shown in the last column. We compare MDF against spectral residual (SR[3]), frequency-tuned saliency (FT [1]), saliency filters (SF [5]), geodesic saliency (GS [6]), hierarchical saliency (HS [7]), regional based contrast (RC [2]), manifold ranking (MR [8]), optimized weighted contrast (wCtr* [9]) and discriminative regional feature integration (DRFI [4]).