Size and Location Matter: A New Baseline for Salient Object Detection

Long Zhao[1], Shuang Liang[1(✉)], Yichen Wei[2], and Jinyuan Jia[1]

[1] Tongji University, Shanghai, China
shuangliang@tongji.edu.cn
[2] Microsoft Research, Beijing, China

Abstract. Recent years have seen many complex models proposed for salient object detection and progressing results. However, less has been done to justify the need for such complex models as there lacks sufficient comparison to simple baselines on more challenging datasets. In this work, we propose a new baseline method for saliency detection. It simply considers a large region close to the image center as salient, and defines the saliency of a region as the product of its size and centerness. As accurate image segmentation problem is difficult by itself, we propose novel techniques that can estimate these attributes using superpixels in a soft manner, without the need to perform hard image segmentation. Our approach is based on very simple concepts and implementation, but already achieves very competitive results, especially on challenging datasets. It is further shown highly complementary with the state-of-the-art. Therefore we believe our method serves as a strong baseline and would enhance the problem understanding for future work.

1 Introduction

Salient object detection has attracted a lot of research interests in recent years [1]. The problem is inherently ambiguous since there lacks common definitions and criteria of "what a salient object is". Consequently, the research in this area presents a great amount of diversity, from low level features to high level methodologies. While many new methods have been proposed and steady improvements in evaluation have been shown, it is still unclear to tell how well and to what extent this problem has been solved.

We observed two issues in the current field: complex methodologies and insufficient evaluation. First, recent works adopt more complex models. The saliency models have evolved from the earlier simple contrast based methods [2–5] and frequency analysis based methods [6,7], to more complex ones such as gaussian mixture appearance models [8], low rank matrix recovery [9], multi-scale segmentation and optimization [10], graph based manifold ranking [11], formulation as a submodular optimization [12], hypergraph modelling [13], Markov Chain [14], learning based [15], and fusion of multiple models [16]. All of these models are well motivated and explained from their own viewpoints, and have been shown working well. However, due to their high complexities and large differences, it is
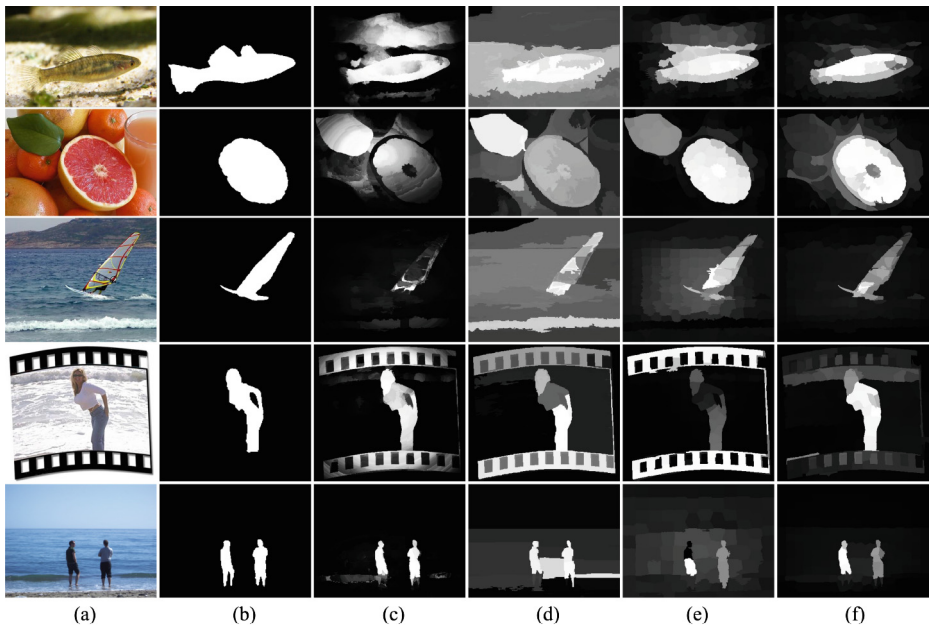
Fig. 1. Saliency detection results on challenging examples. (a) input images; (b) ground truth; (c)–(e) results from the state-of-the-art methods [10,11,19]; (f) our results.

very hard to find how different methods are related and identify what is really working for saliency detection. In other words, it is unclear whether such high complexities are essential or not.

The second issue is that evaluation is mostly performed on the simple ASD [7] or MSRA [2] datasets. It has been well recognized that these datasets are biased to contain a large object near the image center with strong contrast to the background, thus too simple. Although several other more challenging datasets have been proposed, such as SED1 [17], SED2 [17], SOD [18], and ECSSD [10], they are less used in evaluation. While the performance on the simple ASD dataset nowadays is close to saturate, it is relatively unclear whether the good models on ASD can be generalized to more challenging datasets.

This work is a try to address the above two issues by proposing a simple baseline method and showing strong results. Our method just uses two basic concepts: the size and location of a region for determining its saliency. Observing that *larger image regions closer to the image center are more salient*, we define the saliency of a region as the product of its size and centerness. Our definition is intuitive and consistent with human visual perception. The problem is how to compute such concepts reliably.

Region size is clearly informative but has been rarely used before. This is probably because accurate image segmentation problem itself is difficult and there is no good enough segmentation algorithm. While region center has been well known

to be useful for saliency estimation, its usage in previous work is usually overly simple, non-adaptive (such as a gaussian centered on the image) and does not work well for images with different spatial object/background compositions. Our approach is based on a key observation that *geodesic distances between image superpixels essentially encode the segmentation information*. We therefore propose a superpixel based and unified geodesic filtering framework to compute these concepts in a simple and robust manner: (1) it computes approximate region sizes without actually performing image segmentation; (2) it estimates relative region locations with respect to the image center adaptively.

We treat our approach as a baseline because both its concept and implementation is simple, and it can be easily extended or combined with more sophisticated models. Nevertheless, our results are quite strong and encouraging. Extensive experimental comparisons on all above datasets show that our method compares favorably with many recent state-of-the-art complex models. Specifically, it is the best on SED2 [17] and SOD [18], and the second best on SED1 [17]. The examples in Fig. 1 show different challenges for previous methods: low-contrast object (fish, boat), high-contrast but off-center background region (green leaf), complex object/background composition (film), and multiple small objects (beach). Our method works well on such difficult examples while previous methods produce noisy results.

The second encouraging finding is that, after simply combining our results with others, all previous methods are significantly improved and new state-of-the-art results are achieved. Furthermore, the gaps between them before combination are also reduced. This illustrates that these concepts underlying our approach are highly effective and complementary to previous works.

To summarize, this work tackles the saliency detection problem using a basic principle: a large and central region is salient. Our baseline compares favorably and is highly complementary with much more sophisticated models across various datasets. The simplicity, when equipped with strong results, convinces us that the proposed concepts reveal more the essence of saliency detection problem and challenge the necessity of adopting more complex models. Besides the technical contribution, we also expect this work to inspire the field and encourage beneficial changes in mindset.

2   Geodesic Connectivity and Filtering

The geometric attributes such as size and location of an image region are important for determining its saliency. However, extracting good image regions is a challenging problem by itself. All off-the-shelf image segmentation algorithms have the similar problems of how to choose appropriate parameters automatically. Usually, the same parameters could produce different results over different images and this in turn leads to unstable region attributes.

We present simple methods to estimate the size and location of an image region, without actually performing an image segmentation, thus alleviating the

above problems. It performs on a regular superpixel image representation. The parameters are easy-to-set and the results are stable. It is based on a continuous measure of how well any two superpixels are spatially connected, called *geodesic connectivity* in this work. Based on the connectivity measure, we further define a basic operation, called *geodesic filtering*.

An image is first decomposed into a few hundreds of superpixels (200 in our implementation) of similar sizes and regular boundaries, using the recent SLIC algorithm [20]. An undirected weighted graph is created by connecting adjacent superpixels. The edge weight $w_{i,j}$ between superpixels i and j is the Euclidean distance between the average colors of the superpixels in CIELab color space. The geodesic distance, or the length of the shortest path, between any two superpixels $geo\_dist(i,j)$ is defined as

$$geo\_dist(i,j) = \min_{i=v_1,v_2,...,v_n=j} \sum_{k=1}^{n-1} w_{v_k,v_{k+1}} \qquad (1)$$

where $v_1, v_2, ..., v_n$ is a path in the graph linking nodes i and j. Without loss of generality, $geo\_dist(i,i)$ is defined as 0. We then define the *geodesic connectivity* measure as

$$geo\_con(i,j) = exp(-\frac{geo\_dist^2(i,j)}{2\sigma^2}) \qquad (2)$$

The geodesic distance measures the accumulated differences in appearance between two superpixels and the geodesic connectivity characterizes how well they are spatially connected. For the superpixels in the same homogeneous region, the geodesic distance is close to 0 and the connectivity is close to 1. Otherwise, the geodesic distance is large and the connectivity is close to 0. Thus, a superpixel only has large connectivity values for superpixels in the same homogeneous region, and has near zero connectivity values for the other superpixels. Noting this, *the geodesic connectivity measure actually encodes the information of image segmentation in an implicit and soft manner*. It is intuitive, easy-to-implement, and stable. The only important parameter is σ. We found that the performance is stable when $\sigma \in [10, 20]$. It is set to 15 empirically.

We then define a *geodesic filtering* process to measure the properties of image regions from superpixels. Suppose we have a primitive region property map M in superpixel representation, that is, $M(i)$ is the property value of superpixel i, the geodesic filtering computes the property of the region that superpixel i belongs to as

$$\mathcal{GF}(M,i) = \frac{\sum_{j=1}^{N} geo\_con(i,j) \times M(j)}{\sum_{j=1}^{N} geo\_con(i,j)} \qquad (3)$$

where N is the number of superpixels.

Equation (3) is a global filtering of the property map M using geodesic connectivity as weights. It aggregates and smoothes the property values within the same homogeneous region. After filtering, all superpixels in the same region have similar property values of that region. By removing the normalization part (the denominator) in Eq. (3), we obtain an un-normalized version of the filtering,

| (a) | (b) | (c) | (d) | (e) |
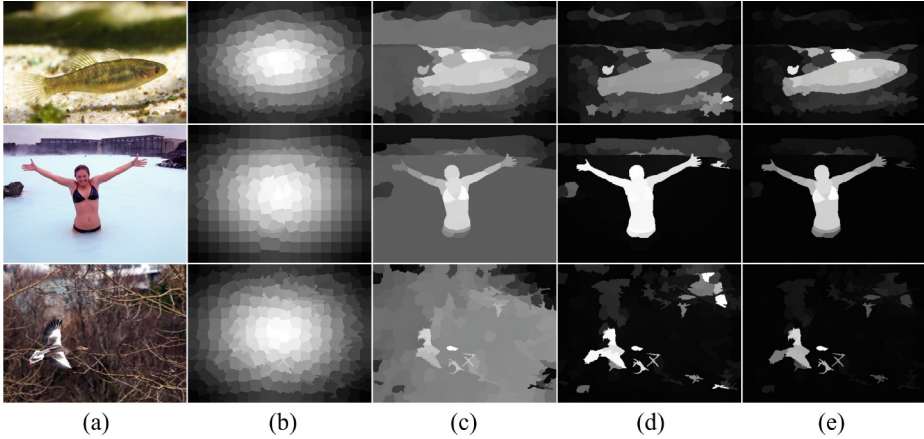
Fig. 2. Illustration of centerness computation. (a) input images; (b) superpixel based gaussian map C_{gau}; (c) geodesic filtered gaussian map SC_{gau} in Eq. (4); (d) image boundary based centerness map C_{bnd} in Eq. (5); (e) our final centerness map C in Eq. (6).

denote as $\tilde{\mathcal{GF}}$. It performs summation instead of averaging. Compared to using a hard image segmentation, our method usually produces smoother and more stable results. The example results before and after geodesic filtering are shown in Figs. 2(b) and (c).

We note that the geodesic saliency propagation approach in [21] shares certain similarity with our work because it essentially applies geodesic filtering to refine an input coarse saliency map. It therefore can be considered as a post-processing and a special case of ours. By contrast, our approach is motivated and derived from a more general viewpoint: we analyze the relation of geodesic distance and segmentation, and generalize the geodesic filtering as a framework to compute more useful region properties (size and centerness) for saliency estimation, which are novel and effective.

3   Our Approach

3.1   Adaptive Computation of Region Centerness

Many saliency methods are biased to assign image center regions with higher saliency. However, previous methods simply use a gaussian fall-off map with mean at the image center and a fixed radius. Such a map does not consider the image content and is problematic for off-center objects or multiple objects. Some methods re-estimate the mean and radius of the gaussian map from an initial saliency map and then refine the saliency map accordingly. This strategy is still not suitable for multiple objects and highly depends on the quality of the initial saliency map.

We propose a simple adaptive method to compute the *centerness* of image regions that alleviates the above problems. We start with a gaussian fall-off map

with mean at image center and standard deviation equals to 10% of the image dimension (the shorter of image width and height). This gaussian map is then turned into a superpixel based version: all the pixels in the same superpixel have their values averaged. We denote the superpixel based gaussian map as C_{gau}. It is exemplified in Fig. 2(b). This map is blocky and uneven in homogeneous image regions. It is then smoothed using geodesic filtering as

$$SC_{gau} = \mathcal{GF}(C_{gau}) \tag{4}$$

The smoothed maps are shown in Fig. 2(c). It is much better but still unsatisfactory because the large background regions usually cover the central parts of the gaussian map and still have large 'centerness' values.

To reduce such errors, we notice that the large background regions also touch the image boundaries. However, special care should be taken because the objects often also do. We further notice that background regions are more widely distributed and more heavily connected to image boundaries than objects: an object seldom touches different sides of the image boundary, while background usually does. We then define a new centerness map C_{bnd} with respect to the four sides of the image boundary, where the value of a superpixel i is computed by considering its geodesic distances to the four sides,

$$C_{bnd}(i) = \sqrt[4]{\mathcal{L}(i) \times \mathcal{T}(i) \times \mathcal{R}(i) \times \mathcal{B}(i)} \tag{5}$$

where $\mathcal{L}(i), \mathcal{T}(i), \mathcal{R}(i)$, and $\mathcal{B}(i)$ are the geodesic distances of superpixel i to the left, top, right, and bottom boundaries, respectively. We add a small constant value to the four distances to avoid the degenerate case when they are equal to 0. Example results of C_{bnd} are shown in Fig. 2(d). The large background regions in Fig. 2(c) are suppressed accordingly.

Our measure in Eq. (5) differs from the work in [11,22] in tricky but important ways. This is illustrated in Fig. 3. The method in [22] simply uses the geodesic distance of a superpixel to the entire image boundary. This is very sensitive for touching-boundary objects, as shown in Fig. 3(b). The method in [11] uses the four boundaries separately in its first stage. However, it does not exploit the concept of geodesic connectivity but uses a complex optimization based on manifold ranking. This usually produces results that are hard to understand, as shown in Fig. 3(c). By contrast, our measure better retains the boundary-touching objects and removes most large backgrounds, as shown in Fig. 3(d).

The two centerness maps in Eqs. (4) and (5) are complementary. Our final centerness map is obtained as the product of the two,

$$C = SC_{gau} \times C_{bnd} \tag{6}$$

Example centerness maps are shown in Fig. 2(e). It is more reasonable than the maps in Figs. 2(c) and (d): the objects in the image center are of higher values and large backgrounds are removed.

Our centerness measure in Eq. (6) is highly adaptive to the image content. It can naturally capture off-center objects and multiple objects, as exemplified in Figs. 1, 2 and 3. This is mainly why our approach outperforms previous methods on images with multiple objects.

| (a) | (b) | (c) | (d) |
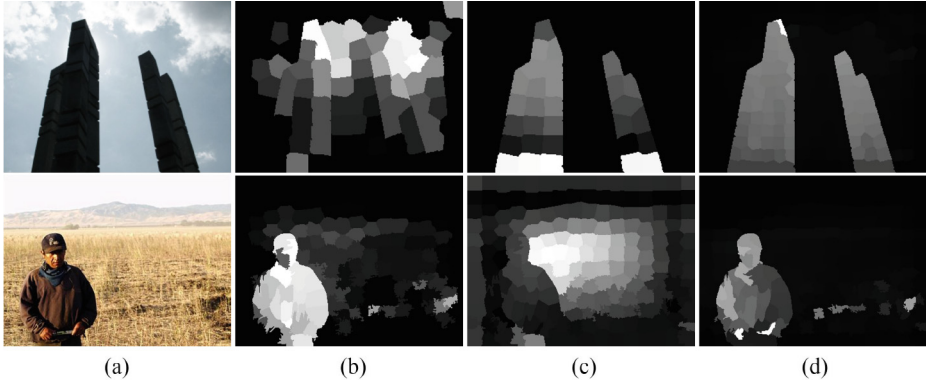
Fig. 3. Illustration of the advantage of our centerness map C_{bnd}. (a) input images; (b) results in [22]; (c) first stage results in [11]; (d) our results of C_{bnd} in Eq. (5).

3.2   Approximate Computation of Region Size

Although the concept of region size is intuitive, it is seldom used in previous work. One possible reason is that it is almost impossible to compute the region size accurately, as image segmentation could be unstable and generate inaccurate regions.

We point out that an accurate segmentation may be unnecessary. Since the superpixels are of similar sizes and shapes, our basic idea is to count the number of superpixels in a homogeneous region and use it as an approximate size of the region. This is done in a soft manner using the geodesic filtering approach in Sect. 2. Let N be the number of superpixels, we denote U as a uniform map that has the same normalized area $\frac{1}{N}$ for all the superpixels. We compute the region size map as

$$A = \tilde{\mathcal{G}}\tilde{\mathcal{F}}(U) \tag{7}$$

Note that we use the un-normalized version of geodesic filtering so for each superpixel it "sums" all superpixels in the same homogeneous region of it, which is the region size. Compared to hard image segmentation methods, our "soft" approach produces more stable and smoother results. This is exemplified in Fig. 4. We tested one of the most widely used image segmentation method in [23]. It has a few parameters. We tried different values and found it is hard to find common parameters that produce reasonable results for different images. We also tried normalized cut and mean shift segmentation algorithms, and found the similar problem. By contrast, our method computes stable and smooth region size maps and does not have the difficult parameter selection problem.

Our final saliency map is simply defined as the product of region size and centerness, as

$$S(i) = C(i) \times \sqrt{A(i)} \tag{8}$$

Note that we use the square root of region size to make the product less sensitive to the region size, which is found useful heuristically.
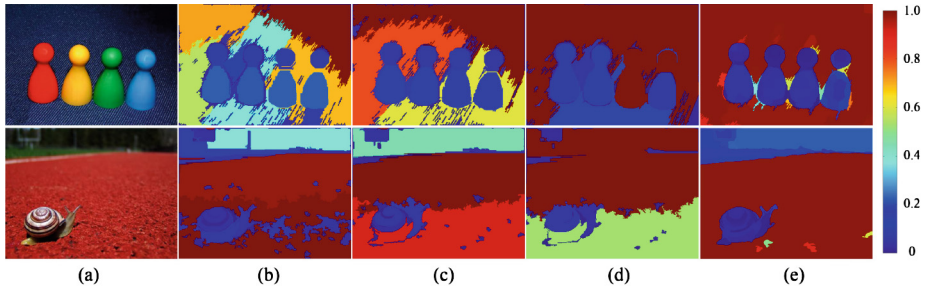
Fig. 4. (Better viewed in color) Example results of computing regions' size using a segmentation method and our method. (a) input images; (b) – (d) region size maps using the segmentation method in [23] with different parameters; (e) region size map of our method. The region size values are normalized to $[0, 1]$ and visualized in color (Color figure online).

4 Experiments

In the experiments, we use six standard benchmark datasets, ASD [7], MSRA [2], SED1 [17], SED2 [17], SOD [18] and ECSSD [10]. ASD [7] and MSRA [2] are relatively simple as there is only one large object near the image center. Note that we obtain the pixel-wise labeling of the MSRA dataset from [15]. The remaining four datasets are more challenging. SED1 [17] and SED2 [17] each contain 100 images with great diversity in object sizes and locations. SOD [18] includes 300 images of complex scenes and multiple objects. It is considered as the most difficult dataset in [1]. ECSSD [10] is a recent dataset extended from CSSD [10]. It includes 1000 images of complex scenes.

We use the standard precision-recall curves (PR curves) and F-measures as evaluation metrics. Given a saliency map, a PR curve is obtained by generating binary masks with a threshold varying from 0 to 255 and comparing these masks against the ground truth. The PR curves are then averaged on each dataset. We follow [7] to compute F measure. For each saliency map, an adaptive threshold (1.5 times of the average saliency) is used to generate a binary mask and precision/recall value. F-measure is then computed as

$$F_\beta = \frac{(1 + \beta^2) \times Precision \times Recall}{\beta^2 \times Precision + Recall} \tag{9}$$

We set $\beta^2 = 0.3$ as in [7] to highlight precision.

We compare with eight recent state-of-the-art methods: saliency filter (SF) [5], geodesic saliency (GS_SP, short for GS) [22], soft image abstraction (SIA) [8], low rank saliency (LRS) [9], hierarchical saliency (HS) [10], dense and sparse reconstruction (DSR) [19], salient region detection by UFO (UFO) [16] and manifold ranking (MR) [11]. There are many other methods in the literature. They are worse than the above methods and not compared for conciseness.

4.1   Comparison with State-of-the-art

Our baseline method is compared with the eight methods. Those methods are also combined with ours, by simply multiplying the two saliency maps. Figure 5 reports the PR curves and F-measures of all methods on all datasets, before and after combining our method.

We can make several interesting observations. Firstly, our method compares favorably with previous works. Besides ASD dataset, our method is always at the top for the other five datasets. Specifically, it is the best on SED2 [17] and SOD [18], and the second best on SED1 [17] in terms of F-measures. We conjecture that this is because other complex methods are more or less over fitted to the simple ASD dataset and do not generalize as well to others. Secondly, after combination all previous methods are significantly improved. The improved results are new state-of-the-art on all datasets. This indicates that our method is highly complementary to previous methods. Especially, SF, GS, SIA and LRS are all improved to a large extent. Lastly, the performance gaps between previous methods are much smaller after combination. For example, while GS and LRS are much worse before combination, they are mostly comparable to the best methods after being improved. Example results of previous methods before and after combining our approach are shown in Fig. 6.

Table 1. The F-measure improvements on different datasets and overall, caused by combining one method to all the other methods, averaged on all other methods. The top two most complementary methods on each dataset are highlighted in bold and underlined bold, respectively.

	ASD [7]	MSRA [2]	SED1 [17]	SED2 [17]	SOD [18]	ECSSD [10]	All
SF [5]	0.0295	−0.0184	−0.1026	0.0537	−0.0582	−0.0686	−0.0274
GS [22]	0.0510	0.0450	0.0237	0.0505	0.0415	0.0356	0.0412
MR [11]	**0.0783**	**0.0693**	**0.0618**	0.0378	0.0528	0.0675	0.0613
DSR [19]	0.0632	0.0659	0.0538	**0.0594**	**0.0653**	**0.0753**	**0.0638**
LRS [9]	0.0377	0.0361	0.0059	0.0307	0.0237	0.0298	0.0273
UFO [16]	0.0675	0.0594	0.0298	0.0347	0.0301	0.0465	0.0447
SIA [8]	0.0422	0.0314	0.0342	0.0153	−0.0054	0.0135	0.0219
HS [10]	0.0654	0.0586	0.0505	0.0508	0.0430	0.0565	0.0541
Ours	**0.0682**	**0.0699**	**0.0697**	**0.0882**	**0.0647**	**0.0742**	**0.0725**

Table 2. Average running time (seconds per image) of different methods, tested on an Intel 3.39GHz Quad-core CPU. For previous methods, we obtained the implementation from the original authors. SIA and HS are in C++ and others are in Matlab.

Ours	SF [5]	GS [22]	MR [11]	DSR [19]	LRS [9]	UFO [16]	SIA [8]	HS [10]
0.260	0.248	0.323	0.825	4.686	12.147	19.209	0.022	0.217
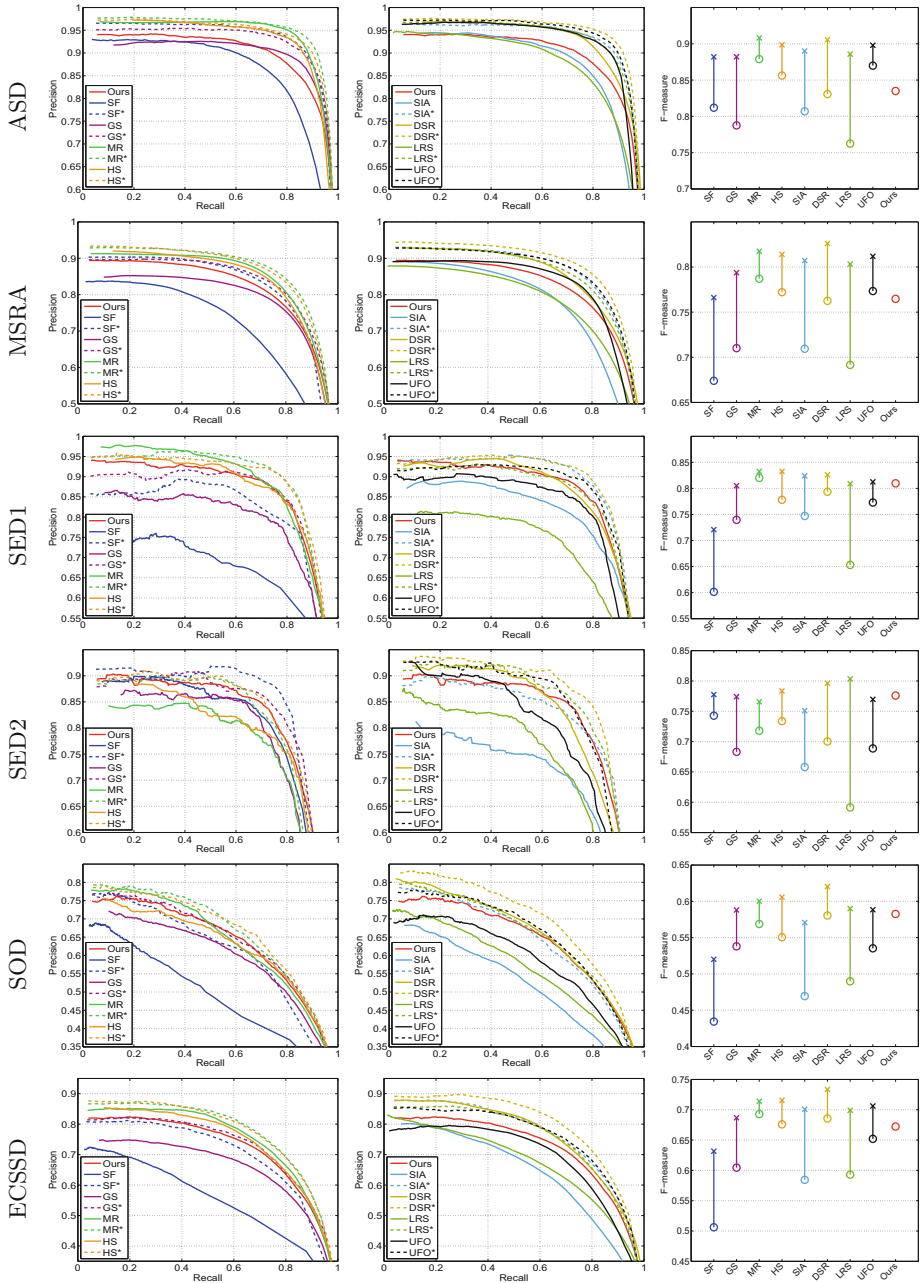
Fig. 5. (Better viewed in color) Precision-recall curves (left, middle) and F-measures (right) of various methods. In the PR curves, results of dotted lines and (*) are obtained by combining our results. In the F-measure, the circle and cross markers are the results before and after combining ours, respectively (Color figure online).

Fig. 6. Example results of eight state-of-the-art methods. For each image, the first row shows the input image and their original results. The second row shows the ground truth and their improved results after combining our approach.
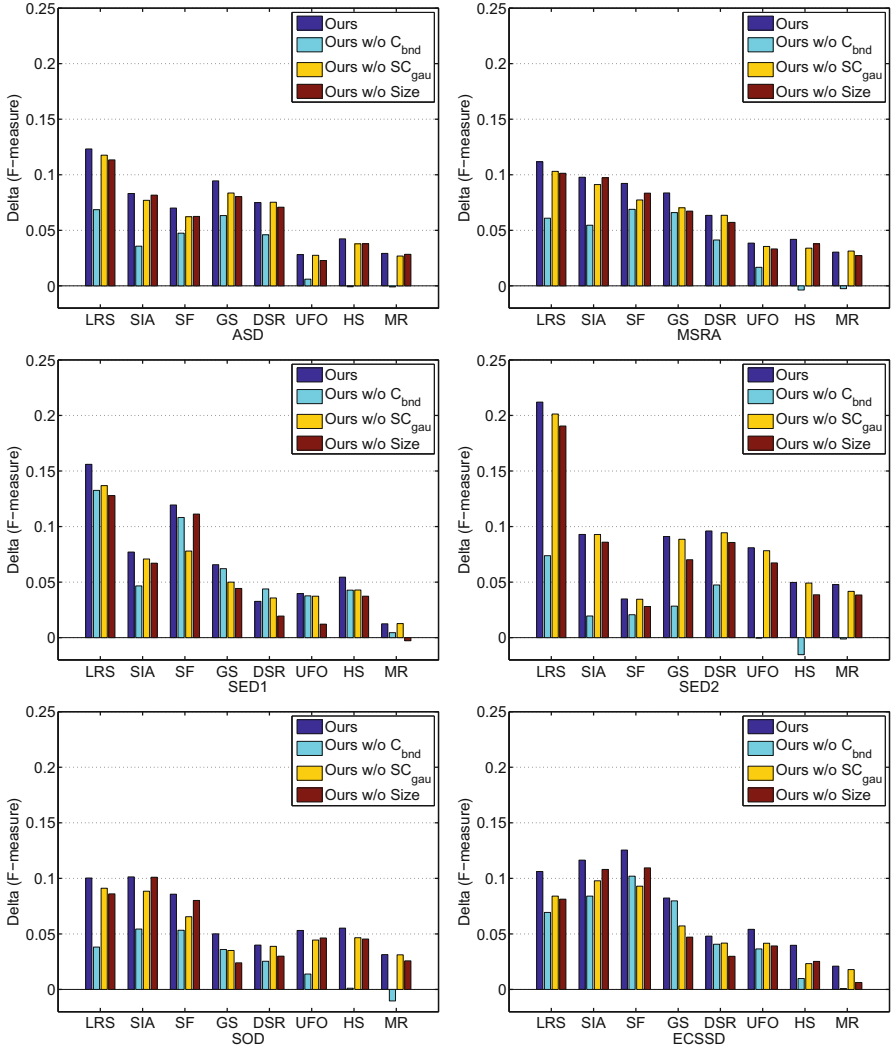
Fig. 7. The relative improvement of F-measure of eight methods caused by our method, our method without using boundary based centerness C_{bnd}, our method without using smoothed gaussian centerness SC_{gau}, and our method without using region size.

We note that it is possible that combining any two good models could produce improvement, as pointed out in [1]. To truly and fairly evaluate how complementary one method is, we report the F-measure improvements by combining it to other methods on the six datasets, averaged on all other methods. Results are shown in Table 1. Indeed, these state-of-the-art methods can improve each other (besides SF), and our method is the most complementary (among the top two
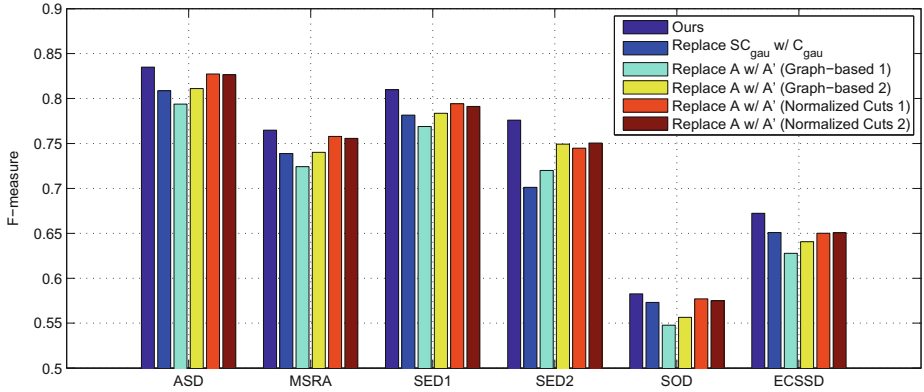
Fig. 8. Evaluation of our region centerness and size by replacing them with other options. See text for details. Graph-based 1 & 2 are computed by [23] with parameters (sigma = 0.5, K = 500, min = 50) and (sigma = 0.5, K = 1000, min = 100), while Normalized Cuts 1 & 2 are computed by [24] with parameters (n = 10) and (n = 20) respectively.

on all datasets, and the best overall), showing that region size and centerness are indeed not well exploited in other methods.

All above results show that our method is highly effective. The running time of all methods are reported in Table 2. Our method is among the fastest ones. Note that our run time includes the superpixel segmentation and shortest path computation in Eq. (1).

4.2   Evaluation of Our Approach

Our result is the product of three components: size map in Eq. (7), the smoothed gaussian centerness map in Eq. (4), and the image boundary based centerness map in Eq. (5). We firstly evaluate their effects by removing each one from the product and checking how much the performance decreases. For conciseness, we only show the relative improvement of F-measure of all the previous methods in Fig. 7. The results demonstrate that the three components all contribute to the improvement and removing any of them would cause performance drop. The results on PR curves are similar.

Evaluation of Geodesic Filtering for Gaussian Centerness Map. To evaluate the effectiveness of applying geodesic filtering in Eq. (4), we remove the filtering and use C_{gau} instead of SC_{gau}, while fixing the other components the same. The results in Fig. 8 show that not using the geodesic filtering clearly decreases the performance.

Evaluation of Region Size. To compare with our 'soft' computation of region size A in Eq. (7), we also compute another region size map A' using hard image segmentation. We segment an image using [23,24], compute the exact size of each region (number of pixels in it), assign each pixel the size of the region enclosing

it, and normalize A' so that its summed value is equal to ours to remove any affection due to magnitude. We then replace A with A' while fixing the other components the same. We test four versions of A': two methods [23,24] and each with two sets of parameters. Results in Fig. 8 show that our soft region size is better and more stable than hard computation of region size, because it is difficult to find good image segmentation parameters for different images.

5   Conclusions

We present a new baseline saliency method. It uses basic principle and concepts of region size and location. We demonstrated how to estimate these attributes with simple techniques, without requiring performing image segmentation. Our method works well across different datasets, including the most challenging ones. It compares favorably with the state-of-the-art and can be easily combined for further improvement. We hope this work can enhance the understanding of salient object detection problem and encourage more works of using simple models that generalize well.

References

1. Borji, A., Sihite, D.N., Itti, L.: Salient object detection: a benchmark. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part II. LNCS, vol. 7573, pp. 414–429. Springer, Heidelberg (2012)
2. Liu, T., Sun, J., Tang, X., Shum, H.Y.: Learning to detect a salient object. In: CVPR (2007)
3. Goferman, S., Manor, L., Tal, A.: Context-aware saliency detection. In: CVPR (2010)
4. Cheng, M., Zhang, G., Mitra, N., Huang, X., Hu, S.: Global contrast based salient region detection. In: CVPR (2011)
5. Perazzi, F., Krahenbuhl, P., Pritch, Y., Hornung, A.: Saliency filters: contrast based filtering for salient region detection. In: CVPR (2012)
6. Hou, X., Zhang, L.: Saliency detection: a spectral residual approach. In: CVPR (2007)
7. Achanta, R., Hemami, S., Estrada, F., Susstrunk, S.: Frequency-tuned salient region detection. In: CVPR (2009)
8. Cheng, M.M., Warrell, J., Lin, W.Y., Zheng, S., Vineet, V., Crook, N.: Efficient salient region detection with soft image abstraction. In: ICCV (2013)
9. Shen, X., Wu, Y.: A unified approach to salient object detection via low rank matrix recovery. In: CVPR (2012)

10. Yan, Q., Xu, L., Shi, J., Jia, J.: Hierarchical saliency detection. In: CVPR (2013)
11. Yang, C., Zhang, L., Lu, H., Ruan, X., Yang, M.H.: Saliency detection via graph-based manifold ranking. In: CVPR (2013)
12. Jiang, Z., Davis, L.S.: Submodular salient region detection. In: CVPR (2013)
13. Li, X., Li, Y., Shen, C., Dick, A., van den Hengel, A.: Contextual hypergraph modelling for salient object detection. In: ICCV (2013)
14. Jiang, B., Zhang, L., Lu, H., Yang, C., Yang, M.H.: Saliency detection via absorbing markov chain. In: ICCV (2013)
15. Jiang, H., Wang, J., Yuan, Z., Wu, Y., Zheng, N., Li, S.: Salient object detection: a discriminative regional feature integration approach. In: CVPR (2013)
16. Jiang, P., Ling, H., Yu, J., Peng, J.: Salient region detection by UFO: uniqueness, focusness and objectness. In: ICCV (2013)
17. Alpert, S., Galun, M., Basri, R., Brandt, A.: Image segmentation by probabilistic bottom-up aggregation and cue integration. In: CVPR (2007)
18. Movahedi, V., Elder, J.H.: Design and perceptual validation of performance measures for salient object segmentation. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 49–56. IEEE (2010)
19. Li, X., Lu, H., Zhang, L., Ruan, X., Yang, M.H.: Saliency detection via dense and sparse reconstruction. In: ICCV (2013)
20. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: Slic superpixels compared to state-of-the-art superpixel methods. PAMI **34**, 2274–2281 (2012)
21. Fu, K., Gong, C., Gu, I., Yang, J.: Geodesic saliency propagation for image salient region detection. In: ICIP (2013)
22. Wei, Y., Wen, F., Zhu, W., Sun, J.: Geodesic saliency using background priors. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part III. LNCS, vol. 7574, pp. 29–42. Springer, Heidelberg (2012)
23. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient graph-based image segmentation. IJCV **59**, 167–181 (2004)
24. Shi, J., Malik, J.: Normalized cuts and image segmentation. PAMI **22**, 888–905 (2000)