

Automatic Salient Object Segmentation Based on Context and Shape Prior

Huaizu Jiang¹
jianghuaizu@gmail.com

Jingdong Wang²
jdwang@microsoft.com

Zejian Yuan¹
yzejian@gmail.com

Tie Liu³
liutie@cn.ibm.com

Nanning Zheng¹
nnzheng@mail.xjtu.edu.cn

Shipeng Li²
spli@microsoft.com

¹ Xi'an Jiaotong University
Xi'an, P. R. China

² Microsoft Research Asia
Beijing, P. R. China

³ IBM China Research Lab
Beijing, P. R. China

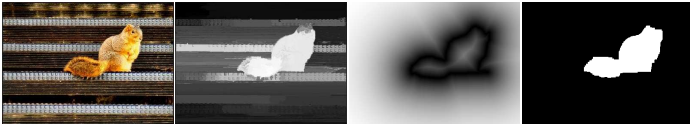


Figure 1: Overview of our proposed salient object segmentation algorithm. (a) input image, (b) context based saliency map, (c) object-level shape prior, (d) result of iterative segmentation.

We propose a novel automatic salient object segmentation algorithm which integrates both bottom-up salient stimuli and object-level shape prior, i.e., a salient object has a well-defined closed boundary. Our approach is formalized as an iterative energy minimization framework.

1 Salient Object Features

Based on our observation, we introduce three characteristics to define a salient object:

1. The salient object is always different from its surrounding context.
2. The salient object in an image is most probably placed near the center of the image.
3. A salient object has a well-defined closed boundary.

1.1 Context-based Saliency Computation

Our saliency is defined based on multi-scale superpixels, which are generated by fragmenting the image using N groups of different parameters [2]. According to characteristic 1, a region is salient if it is distinguished from its immediate context, defined as a set of spatial neighbors in our scenario. The difference between regions is measured by area weighted color histogram distance. And according to characteristic 2, we introduce Gaussian falloff weight to highlight regions near the image center. Finally, we propagate saliency value from multiple regions to pixels, weighted by the color distance between pixels and region centers, to get saliency map S_m . Such multi-superpixel-scale enhancement makes our saliency map more robust in cluttered background.¹

1.2 Shape Prior Extraction

According to characteristic 3, a salient object has a closed boundary. We first construct an edge map E . The edge map consists of a set of line segments which are obtained from an edge detector, followed by a line fitting step. We refer to straight line segments as *detected segments*. Our shape prior extraction can then be formalized to find an optimal closed contour by identifying a subset of detected segments in E and connecting them together. The optimal closed contour can be defined as a ratio form balancing the gaps along the contour and saliency value of pixels located inside the contour. Such an optimal contour can be found efficiently in polynomial time using the ratio contour algorithm [4]. Shape prior S_p is finally defined as the distance transform of such a salient contour.

2 Salient Object Segmentation Framework

2.1 Energy Model for Salient Object Segmentation

Given input image I , saliency map S_m (section 1.1), and shape prior S_p (section 1.2), our goal is to find the label set L , where $l_p \in \{0, 1\}$ for each pixel p , 0 for background and 1 for salient object (foreground). Salient object segmentation can be achieved through energy minimization, which

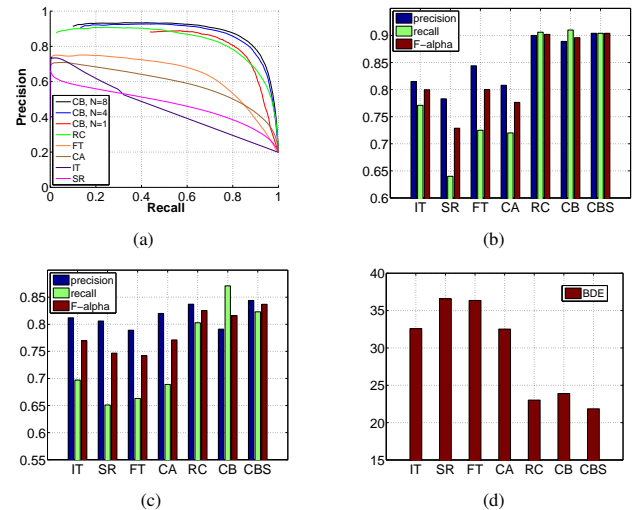


Figure 2: Quantitative comparison of different methods.

combines saliency with object boundary information. Different from previous works, we incorporate shape prior into such a model to better segment the salient object.

2.2 Iterative Energy Minimization

The initial saliency map and shape prior are only rough estimation of the salient object. After binary segmentation, both of them can be re-estimated more accurately. We first update the saliency map by comparing the color histograms of salient object (foreground) and background. Based on the new saliency map, we can update the shape prior, and then re-segment the image. We run iterative energy minimization until convergence. The algorithm of our iterative segmentation is summarized below:

Algorithm 1 $L = \text{SalientObjectSegmentation}(I)$

- 1: Calculate saliency map S_m .
- 2: Extract shape prior S_p based on S_m .
- 3: Segment image by energy minimization.
- 4: Update saliency map S_m based on current segmentation L .
- 5: Go to step 2 to update shape prior S_p , then re-segment image until convergence.

3 Experimental Results

Quantitative comparisons of different methods including IT, SR, FT, CA and RC (see full paper for details) on two benchmark datasets [1, 3] demonstrate our proposed algorithm CBS, which integrates bottom-up saliency and object-level shape prior, outperforms state-of-the-art methods.

- [1] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-tuned salient region detection. *IEEE CVPR*, 0:1597–1604, 2009.
- [2] Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2):167–181, 2004.
- [3] Tie Liu, Jian Sun, Nan-Ning Zheng, Xiaou Tang, and Heung-Yeung Shum. Learning to detect a salient object. *CVPR*, 0:1–8, 2007.
- [4] Song Wang, Toshiro Kubota, Jeffrey Mark Siskind, and Jun Wang. Salient closed boundary extraction with ratio contour. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(4):546–561, 2005.

¹code is available at <https://sites.google.com/site/jianghz88/>