PROBABILISTIC SALIENT OBJECT CONTOUR DETECTION BASED ON SUPERPIXELS

Huaizu Jiang[1] *Yang Wu*[2] *Zejian Yuan*[1]

[1]Xi'an Jiaotong University, Xi'an, China
[2]Kyoto University, Kyoto, Japan

ABSTRACT

In this paper, we propose a data-driven approach to detect the probabilistic salient object contour, which is formulated as predicting the probability of superpixel boundaries being on the object contour based on the learned regressor. Each superpixel boundary is jointly described by the superpixel saliency, superpixel contrast, and boundary geometry features. Experimental results on the benchmark data set validate the effectiveness of our approach. Furthermore, we demonstrate that the predicted probabilistic salient object contour is useful for improving the multiple segmentations for salient object detection.

Index Terms— salient object contour, superpixels

1. INTRODUCTION

Though many computational models exist for either detecting the salient object or predicting human beings' eye fixations, few of them focus on the salient object contour detection. In this paper, we propose a data-driven approach to give a probabilistic prediction of the salient object contour based on superpixels. Our basic assumption is that the boundary of the salient object aligns well with superpixel boundaries. By segmenting the input image into superpixels, we cast the problem to estimating the probability of each superpixel boundary being on the object contour. Illustration of our approach is shown in Fig. 1.

To describe each superpixel boundary, we propose a set of features including superpixel saliency, superpixel contrast, and boundary geometry. The feature vector will be mapped to a probability score based on the learned regressor. Our data-driven approach can automatically integrate all the features and pick up the most discriminative ones. Additionally, detecting the salient object contour on the superpixels rather than on pixels allows us to consider more complicated features such as color and texture histograms.

We train and test our model on the benchmark MSRA-B data set [1] to validate the effectiveness of our approach. Furthermore, we show that our probabilistic salient object contour detection is useful for improving multiple segmentations for salient object detection [2].

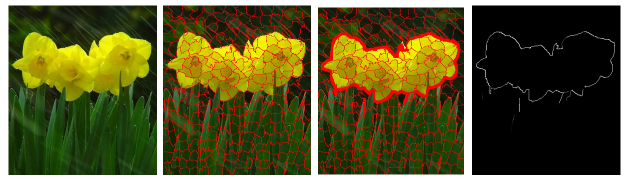


Fig. 1. Illustration of our approach. **Left:** input image. **Middle Left:** superpixels. **Middle Right:** illustration of the probabilistic salient object contour. The wider the superpixel boundary, the more salient it is. **Right:** probabilistic salient object contour detection result.

1.1. Related Work

Contour detection is a classical problem in computer vision. And recently great advances have been reported in [3] which integrates various multi-scale local cues into a globalization framework based on spectral clustering. The local cues are computed by applying oriented gradient operators at every pixels in the image. Similar to [3], we also give a probabilistic prediction of the contour of an object. But compared with such a pixel-based method, our superpixel-based approach considers more complicated features like color and texture histograms.

In [4], the salient contour is computed by connecting the detected edge segments, where the edge connecting problem is formulated as optimizing a ratio form and can be efficiently solved in polynomial time. And recently Kennedy *et al.* [5] propose to detect the salient contour on a weighted graph in which each node corresponds to a detected contour segment. The contour detection problem is then defined as finding a cycle on the graph by solving a Hermitian eigenvalue problem. Slightly different from these approaches to detect the contour of a general object, our method focuses on finding the contour of a salient object by considering the visual saliency information.

There also exist approaches for detecting the edges of a salient object, *e.g.* [6], which integrates local and regional features on multiple scales. The goal of [6] is to detect the detailed *edges* of the salient object, however, and we try to give a probabilistic estimation of the *contour*.

Table 1. Superpixel features adopted to predict the probabilistic salient object contour. The χ^2 distance between histograms $\mathbf{h}_1$ and $\mathbf{h}_2$ is defined as $\chi^2(\mathbf{h}_1, \mathbf{h}_2) = \sum_{i=1}^{b} \frac{2(h_{1i}-h_{2i})^2}{h_{1i}+h_{2i}}$ with b being the number of histogram bins. Abbreviations: absolute (abs.), difference (diff.).

Superpixel Boundary Features			
Superpixel Saliency	**Dim**	**Superpixel Contrast**	**Dim**
s1. the abs. diff. of average RGB values w.r.t. $\mathcal{B}$	3	c1. the abs. diff. of average RGB values	3
s2. the χ^2 distance of RGB histogram w.r.t. $\mathcal{B}$	1	c2. the χ^2 distance of RGB histogram	1
s3. the abs. diff. of average HSV values w.r.t. $\mathcal{B}$	3	c3. the abs. diff. of average HSV values	3
s4. the χ^2 distance of HSV histogram w.r.t. $\mathcal{B}$	1	c4. the χ^2 distance of HSV histogram	1
s5. the abs. diff. of average L*a*b* values w.r.t. $\mathcal{B}$	3	c5. the abs. diff. of average L*a*b* values	3
s6. the χ^2 distance of L*a*b* histogram w.r.t. $\mathcal{B}$	1	c6. the χ^2 distance of L*a*b* histogram	1
s7. the abs. diff. of average responses of filter bank w.r.t. $\mathcal{B}$	15	c7. the abs. diff. of average responses of filter bank	15
s8. the χ^2 distance of maximum response of filter bank w.r.t. $\mathcal{B}$	1	c8. the χ^2 distance of maximum response of filter bank	1
s9. the χ^2 distance of texton histogram w.r.t. $\mathcal{B}$	1	c9. the χ^2 distance of texton histogram	1
s10. the average x coordinates	1	**Boundary Geometry**	**Dim**
s11. the average y coordinates	1	g1. the average x coordinates	1
s12. the $10th$ percentile of x coordinates	1	g2. the average y coordinates	1
s13. the $10th$ percentile of y coordinates	1	g3. the $10th$ percentile of x coordinates	1
s14. the $90th$ percentile of x coordinates	1	g4. the $10th$ percentile of y coordinates	1
s15. the $90th$ percentile of y coordinates	1	g5. the $90th$ percentile of x coordinates	1
s16. the normalized area	1	g6. the $90th$ percentile of y coordinates	1
		g7. the normalized length	1

The similar idea of detecting the closed contour of the object based on superpixels has been exploited in [7], which formulates the contour detection as a binary labeling problem of each superpixel. Boundaries of the superpixels which are given the positive labels are gathered to form the closed contour of the object. Our approach distinguishes from [7] in the motivation. We are aiming at predicting the saliency probability of each superpixel boundary while [7] focuses on the binary labeling of the superpixels.

2. OUR APPROACH

The input image I will be first segmented into superpixels $\{s_i\}_{i=1}^{N}$. Based on a learned regressor, we then try to predict the probability of each superpixel boundary b_{ij}, which is formed by the adjacent superpixels s_i and s_j, being on the salient object contour. The superpixel boundary b_{ij} is jointly described by the superpixel contrast between s_i and s_j, saliency features of s_i and s_j, and the geometry features of b_{ij}.

2.1. Superpixel Boundary Features

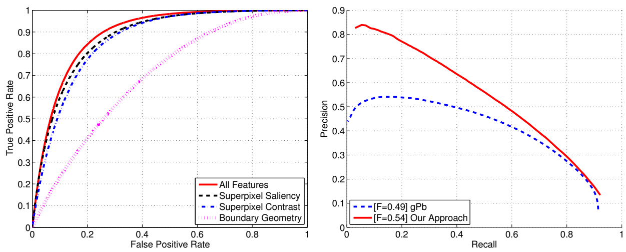
Superpixel Contrast (29 features). If s_i and s_j share the same appearance, most probably they are both from the salient object or the background. Therefore, there is weak evidence that the boundary b_{ij} is on the salient object contour.

We describe the appearance contrast of superpixels from color and texture perspectives. The color contrast of two superpixels are defined as the absolute color difference and the χ^2 distances color histograms in different color spaces for robustness purpose. We apply the filter bank, including 15 filters, proposed in [8] to describe the texture of a superpixel. And the texture contrast of two superpixels include the absolute difference of filter response of each filter and the χ^2 distance of the histogram of maximum responses. We additionally consider the χ^2 distance of the texton [3] histograms of two superpixels.

Superpixel Saliency (2×36 features). If both s_i and s_j are from the cluttered background, though their appearance contrast is large, b_{ij} is actually not on the salient object contour. This indicates that only the superpixel contrast features are not discriminative enough to describe the superpixel boundary. We thus consider the visual saliency features of s_i and s_j. If both of them are visually salient or non-salient, b_{ij} is probably not on the salient object boundary.

To attract human beings' attention, the salient object is most probably placed near the center of the image. In another word, the background is always far away from the image center. Therefore, we define the pseudo-background $\mathcal{B}$ as the 15-pixel wide region of the image border. If the appearance (color and texture) of the superpixel s_i (or s_j) is distinct from $\mathcal{B}$, most probably it is salient. Consequently, we consider the color and texture contrast between s_i (or s_j) and the $\mathcal{B}$. Furthermore, we consider the geometry features of a superpixel, *e.g.*, the superpixel near the image center are more likely to be salient than the one far away from the center.

Boundary Geometry (7 features). The geometry features of b_{ij} may influence its probability of being on the salient contour as well. For example, if b_{ij} is near the image border (and thus far away from the image center), it is less likely to



(a) (b)

Fig. 2. Quantitative evaluation of our approach. (a) importance of superpixel boundary features, and (b) comparison of gPb [3] and our approach on the MSRA-B [1] data set in terms of the precision-recall curve and F-measure.

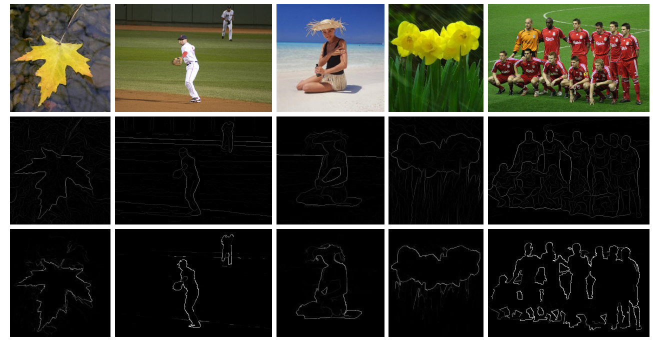


Fig. 3. Visual comparison of different approaches. From top to bottom: input images, probabilistic salient object contour detection using gPb [3] algorithm, and our approach.

be on the salient object contour. Specifically, we consider the average x and y coordinates, 10th and 90th percentile of the x and y coordinates, and the normalized length of b_{ij}.

In summary, we extract a 108-dimension feature vector to describe each superpixel boundary. Detailed descriptions of the features can be found in Table 1.

2.2. Learning

We learn a Boosted Decision Tree (BDT) classifier from a set of training examples to map the feature vector of each superpixel boundary to a scalar, which will be transformed to a probability score using the logistic regression. The training examples include the a set of superpixel boundaries $\{x_1, x_2, \cdots, x_M\}$ and corresponding contour scores $\{y_1, y_2, \cdots, y_M\}$. y_i is set to 0 indicating x_i is not on the salient object contour. Otherwise y_i is set to 1. x_i is said to be on the salient object contour if the two superpixels forming x_i are both from the salient object or the background.

3. EXPERIMENTAL RESULTS

3.1. Setup

We train and test our model on the MSRA-B data set [1]. This data set includes 5000 images, originally containing labeled rectangles. There is a large variation among images such as natural scenes, animals, indoor, and outdoor. We manually segmented the salient object (contour) within the user-drawn rectangle to obtain binary masks. We randomly choose 3000 images for training, and the rest of 2000 images for testing.

For quantitative evaluation, we report the precision-recall curve and F-measure as in [3]. And we adopt the SLIC algorithm [9] to generate the superpixels where the superpixel number is set to 300.

3.2. Importance of Superpixel Boundary Features

To evaluate the importance of our proposed superpixel boundary features, we train BDT classifiers on each feature set, and report the ROC curves on the testing set in Fig. 2 (a). As we can see, superpixel saliency is the most discriminative set. The reason might be that superpixel saliency is useful for suppressing the superpixel boundaries with large superpixel contrast in the cluttered background (or the salient object) where the adjacent superpixels share the same visual saliency. Superpixel contrast performs well since it may find the superpixel boundaries on the salient object contour with large appearance contrast. Since the boundary geometry feature set contains only 9 features, it lacks of enough discriminative power and thus performs worst. Finally, tt is worth noting that when jointly considering all the features, the best performance can be achieved.

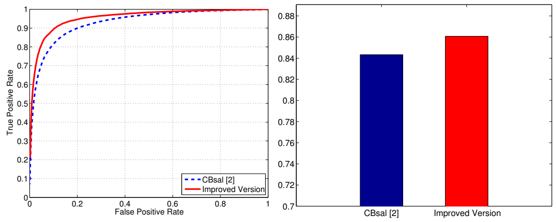
3.3. Salient Object Contour Detection

To validate the effectiveness of our proposed approach, we compare it with state-of-the-art contour detection algorithm [3] on MSRA-B data set. The precision-recall curve along with the F-measure is plotted in Fig. 2 (b). As can be seen, our approach can achieve better performance to detect the salient object contour. It improves by around 10% in terms of the F-measure score compared with gPb [3].

Additional visual comparisons are shown in 3. Compared with gPb, our approach can generate visually better results. For example, our method can to some extent suppress the noisy textures of the background in the first column since the visual saliency information is taken into consideration.

4. IMPROVING MULTIPLE SEGMENTATIONS FOR SALIENT OBJECT DETECTION

In [2], a saliency map is computed based on multiple segmentations which are generated by fragmenting the image by



(a) (b)

Fig. 4. Quantitative comparison of the approach described in [2] (CBsal) and our improved version in terms of (a) ROC curve and (b) AUC (Area Under the ROC Curve) scores.

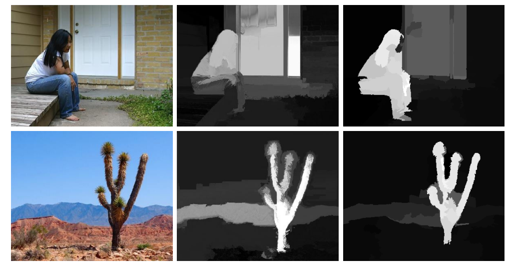


Fig. 5. Visual comparisons of the approach described in [2] (middle) and our improved version (right).

several groups of parameters. Here, we improve the unsupervised multiple segmentations by utilizing the probability of being on the salient object contour for each superpixel boundary.

In specific, an undirected graph $G(V, E)$ is constructed, where the node v_i is corresponding to the superpixel s_i. And v_i and v_j will be connected if s_i and s_j are adjacent. The weight w_{ij} is defined as $1 - p_{ij}$, where p_{ij} is the learned probability for the superpixel boundary $_{ij}$ being on the salient object contour. We apply the graph-based segmentation algorithm [10] on graph G. By varying the parameter k which controls the tolerance of small regions (see the details in [10]), we can generate multiple segmentations.

Compared with the unsupervised multiple segmentations adopted in [2], our improved version can reliably generate coherent multiple segmentations where the superpixels belonging to the salient object (or the background) are more likely to be grouped together. For comparison, we compute the saliency map on the 2000 testing images of MSRA-B data set, where eight segmentations are generated for each image as in [2]. The ROC curve and AUC (Area Under ROC Curve) scores are plotted in Fig. 4. Additional visual comparisons are provided in Fig. 5. As can be seen, much better performance of salient object detection can be achieved by utilizing the salient object contour detection results.

5. CONCLUSION

In this paper, we propose a data-driven approach to predict the probabilistic salient object contour. By mapping the feature vector of each superpixel boundary, we can get its probability being on the salient object contour. Experimental results on MSRA-B data set validates effectiveness of our proposed approach. Additionally, we demonstrate that the probabilistic salient object contour may improve the multiple segmentations for salient object detection [2].

Acknowledgements This work was supported in part by the National Basic Research Program of China under Grant No. 2012CB316400, and the National Natural Science Foundation of China under Grant No. 91120006.

References

[1] Tie Liu, Zejian Yuan, Jian Sun, Jingdong Wang, Nanning Zheng, Xiaoou Tang, and Heung-Yeung Shum, "Learning to detect a salient object," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 2, pp. 353–367, 2011.

[2] Huaizu Jiang, Jingdong Wang, Zejian Yuan, Tie Liu, Nanning Zheng, and Shipeng Li, "Automatic salient object segmentation based on context and shape prior," in *British Machine Vision Conference (BMVC)*, 2011.

[3] Pablo Arbelaez, Michael Maire, Charless Fowlkes, and Jitendra Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, 2011.

[4] Song Wang, Toshiro Kubota, Jeffrey Mark Siskind, and Jun Wang, "Salient closed boundary extraction with ratio contour.," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 4, pp. 546–561, 2005.

[5] Ryan Kennedy, Jean H. Gallier, and Jianbo Shi, "Contour cut: Identifying salient contours in images by solving a hermitian eigenvalue problem," in *CVPR*, 2011, pp. 2065–2072.

[6] M. Holtzman-Gazit, L. Zelnik-Manor, and I. Yavne, "Salient edges: A multi scale approach," in *ECCV Workshop on Vision for Cognitive Tasks*, 2010.

[7] Alex Levinshtein, Cristian Sminchisescu, and Sven J. Dickinson, "Optimal contour closure by superpixel grouping," in *ECCV (2)*, 2010, pp. 480–493.

[8] Thomas K. Leung and Jitendra Malik, "Representing and recognizing the visual appearance of materials using three-dimensional textons," *International Journal of Computer Vision*, vol. 43, no. 1, pp. 29–44, 2001.

[9] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurélien Lucchi, Pascal Fua, and Sabine Süsstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, 2012.

[10] Pedro F. Felzenszwalb and Daniel P. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004.