Automatic Scribble Simulation for Interactive Image Segmentation Evaluation

Bingjie Jiang, Tongwei Ren*, and Jia Bei

[1] State Key Laboratory for Novel Software Technology, Nanjing University, China
[2] Software Institute, Nanjing University, China
bjie.jiang@gmail.com, rentw@nju.edu.cn, beijia@software.nju.edu.cn

Abstract. To provide comprehensive evaluation of interactive image segmentation algorithms, we propose an automatic scribble simulation approach. We first analyze the variety of scribbles labelled by different users and its influence on segmentation result. Then, we describe the consistency and inconsistency of scribbles with normal distribution on superpixel level and superpixel group level, and analyze the effect of connection in scribble for interactive segmentation evaluation. Based on the above analysis, we simulate scribbles on foreground and background respectively by randomly selecting superpixel groups and superpixels with the previously determined coverage values. The experimental results show that the scribbles simulated by the proposed approach can obtain similar evaluation results to manually labelled scribbles and avoid serious deviation in precision and recall evaluation.

Keywords: scribble simulation, interactive image segmentation evaluation, scribble variety, superpixel group

1 Introduction

As the foundation of numerous multimedia applications, interactive image segmentation has been widely utilized in object recognition [1], image retrieval [2] and annotation [3], scene understanding [4], visual tracking [5], social media mining [6], surveillance analysis [7] and so on. It can effectively extract the desired objects from images with the assistance of manual labels, which is used to approximately outline the regions of objects [8]. There have been several types of manual labels applied in the existing interactive image segmentation algorithms, including triple map, boundary box and scribble, in which scribble is commonly used for its simplicity and flexibility in labelling [9].

For different users may provide various scribbles in labelling, an effective interactive image segmentation algorithm should be robust to scribbles, i.e., its performance should not be obviously influenced by the difference of scribbles labelled by different users. It requires to evaluate interactive image segmentation algorithms on the scribbles provided by numbers of users with sufficient variety. However, in the evaluations of the existing algorithms [8, 9], only the scribbles provided by one specific user are used to reduce human labor. Obviously, such evaluations cannot provide a comprehensive comparison for they contain high randomness in scribble labelling and segmentation. Moreover, if the scribbles are intentionally selected, the evaluations may have a bias to some algorithms, which influences the fairness of the evaluations.

To overcome the above problem, we propose an automatic scribble simulation approach for interactive image segmentation evaluation. Based on the ground truths of segmentation results, a number of scribbles with sufficient variety are automatically generated to simulate the labelling results provided by different users, which can provide a comprehensive evaluation of interactive image segmentation algorithms.

The rest of the paper is organized as follows: In section 2, we briefly review the existing interactive image segmentation algorithms and evaluation strategies. In section 3, we introduce the data set used in our experiments. In section 4, we analyze the variety of scribbles labelled by different users and its influence on segmentation results. In section 5, we present the details of our scribble simulation approach. In section 6, the proposed scribble simulation approach is validated by comparing manual labelling and other two automatic simulation approaches. Finally, the paper is concluded in section 7.

2 Related Work

Interactive image segmentation. Amounts of interactive image segmentation algorithms have been proposed in decades. As one of the most representative algorithms, Graph cuts [8] converts each image to a graph and formulates image segmentation as a min-cut energy minimization problem. Grabcut [10] allows to label a rectangle around foreground and improves segmentation performance by iteration strategy. Random walker [11] assigns each unlabelled pixel a maximal probability that a random walker could reach it starting from the existing labels. Geodesic distance [12] uses star-convexity prior and replaces Euclidean rays with geodesic path to exploit the structure of shortest paths. Recently, some researchers extend interactive segmentation from monocular image to other media types, such as binocular image [9], RGB-D image [13] and video [14].

Interactive image segmentation evaluation. In evaluation of interactive image segmentation algorithms, a key problem is how to effectively generate sufficient user labels. Manual labelling with several participants has high labor cost and time consumption even with a facilitate tool [15], and the labelled scribbles cannot be utilized to handle new test images [16]. To overcome the limitations of manual labelling, automatic interactive image segmentation evaluation is studied as in image and video compression [17], resizing [18] and summarization [19]. Some existing approaches attempt to generate the scribbles similar to manual labelling results in appearance [20–22], for example, extracting the sketches of foreground objects and labelling the background around the objects. However, the appearance of scribbles generated by these approaches are constricted by pre-defined rules, which reduces the variety of the generated scribbles. To increase representation flexibility, another approaches use pixel sets instead of connected curves in representing scribbles. Nevertheless, the existing approaches usually sample the pixels by manual defined strategies without considering the characteristics of manual labelling [23].

3 Data Set

We construct our data set with 96 images from Berkeley Segmentation Dataset [16]. Each image contains at least one obvious object, which could be unambiguously explained to users. And these images are also representative of some major challenges of image segmentation, including fuzzy boundary, complex texture and complex lighting conditions. The ground truths of segmentation results are precisely hand-labelled for each image to avoid biases.

To analyze the rules for scribble simulation, we invite five users to manually label the images with The K-Space Segmentation Tool Set [15]. All the users are the students with basic computer operating skills but limited knowledge about interactive image segmentation. Each user is given a clear guidance and enough time to familiarize themselves with the labelling software, and all the labelling operations are carried out by mouse.

4 Analysis of Scribble Variety

4.1 Scribble difference

An instinctive observation is that different users cannot keep high consistency in labelling images with scribbles. In order to validate the observation, we analyze the variety of scribbles labelled by different users. To facilitate the following description, we indicate the five users with A, B, C, D and E. To the kth image, the scribbles labelled by user n are represented as s_n^k, here $n \in \{A, B, C, D, E\}$. And the scribbles labelled by user n on all the images are represented as S_n.

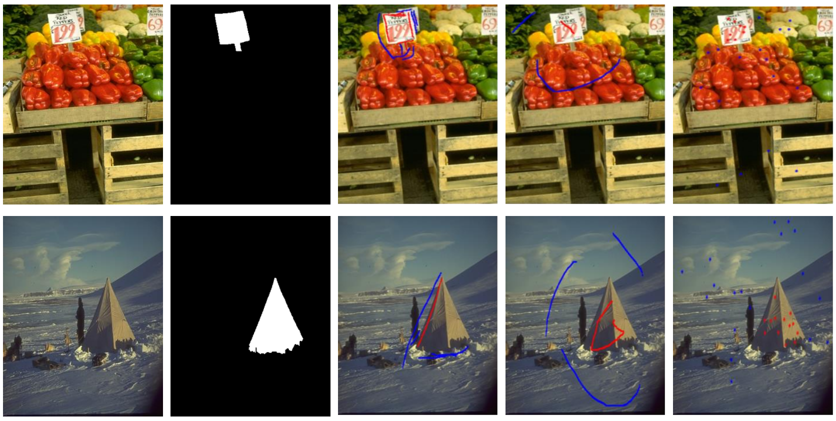
We first analyze the difference of scribbles by pair-wise intersection rate. To the kth image, the intersection rate of the scribbles labelled by user m and n is calculated as $\phi_{m,n}^k = (s_m^k \cap s_n^k)/(s_m^k \cup s_n^k)$. And the average intersection rate of all the scribbles labelled by user m and n is calculated as $\bar{\phi}_{m,n} = \frac{1}{K}\sum_{k=1}^{K} \phi_{m,n}^k$, here $K = 96$ is the number of images in our data set. Table 1 shows the average interaction rates between the scribbles labelled by all the users on foreground and background, respectively. It is observed that the values of average intersection rates between the scribbles labelled by different users are quite low.

4.2 Influence on segmentation result

We further evaluate the influence of different scribbles on segmentation results. We utilize the scribbles labelled by different users as the inputs of interactive image segmentation algorithms. In our experiments, four interactive image segmentation algorithms, including graph cuts (GC) [8], geodesic star convexity (GSC) [12], random walker (RW) [11], and geodesic shortest path (GSP) [14],

Table 1. Pixel level average intersection rates of different scribbles on foreground and background.

foreground						background					
	S_A	S_B	S_C	S_D	S_E		S_A	S_B	S_C	S_D	S_E
S_A	–	2.4%	3.5%	2.0%	1.7%	S_A	–	0.9%	0.9%	0.1%	0.6%
S_B	2.4%	–	2.7%	2.2%	2.3%	S_B	0.9%	–	1.0%	0.2%	0.7%
S_C	3.5%	2.7%	–	2.1%	2.0%	S_C	0.9%	1.0%	–	0.2%	0.7%
S_D	2.0%	2.2%	2.1%	–	1.8%	S_D	0.1%	0.2%	0.2%	–	0.7%
S_E	1.7%	2.3%	2.0%	1.8%	–	S_E	0.6%	0.7%	0.7%	0.7%	–



(a) (b) (c) (d) (e)

Fig. 1. Examples of manually labelled scribbles and automatic simulation results. (a) Original image. (b) Ground truth of segmentation result. (c)-(d) Scribbles labelled by different users. (e) Automatic simulation result.

are used, whose implementations are all provided in [12]. Totally, $96 \times 5 \times 4$ segmentation results are generated by the four algorithms initialized with all the scribbles. To the five segmentation results generated by one algorithm on each original image, we calculate the percentages of pixels which occur as foreground in one, two, three, four or five segmentation results, respectively.

Fig. 2 shows the boxplots of pixel co-occurrence percentages in different numbers of segmentation results for each segmentation algorithm. We can observe that the pixel co-occurrence percentages decline greatly when the numbers of segmentation results increase for all the algorithms. It means the segmentation results generated by the scribbles from different users are quite inconsistent. Moreover, to some segmentation algorithms which can generated relatively higher consistent results, such as GC and GSC, we can also find that their segmentation results contain high variability (larger lengthes of the boxes in Fig. 2 (a) and (b)). Therefore, multiple scribbles with sufficient variety is required for comprehensively evaluating the performance of interactive segmentation algorithms.

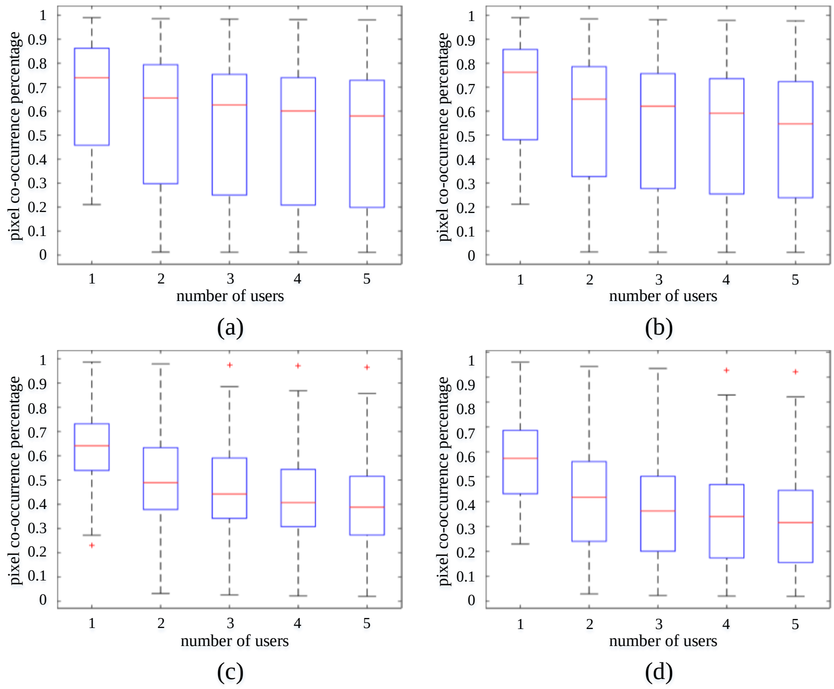


Fig. 2. Percentages of pixel co-occurrence as foreground in different numbers of segmentation results. (a) GC. (b) GSC. (c) RW. (d) GSP.

5 Automatic Scribble Simulation

5.1 Scribble consistency on superpixel and superpixel group levels

For the scribbles labelled by different users have high inconsistence on pixel level, we analyze the scribbles on superpixel levels for expecting high consistency. we use the simple linear iterative clustering algorithm [24], which is implemented by VLFeat open source library [25], to cluster image pixels into compact and nearly uniform superpixels. To each superpixel, we consider it to be labelled by a scribble if one or more pixels within it are labelled by the scribble. Similar to pixel-level scribble consistency analysis, we analyze the scribble consistency on superpixel level by calculating the pair-wise intersection rates of the scribbles labelled by different users.

Table 2 shows the intersection rates of the scribbles labelled by different users on superpixel level. Here, $\widehat{S}_n$ denotes the scribbles labelled by user n on all the images on superpixel level. Compared to pixel level analysis in Table 1, the scribbles have higher consistency on superpixel level than on pixel level, but the values of intersection rates are not high enough especially on background.

To further explore the consistency of different scribbles, we divide the superpixels into groups by quantifying them on RGB color space. We uniformly

Table 2. Superpixel level average intersection rates of different scribbles on foreground and background.

foreground						background					
	$\widehat{S}_A$	$\widehat{S}_B$	$\widehat{S}_C$	$\widehat{S}_D$	$\widehat{S}_E$		$\widehat{S}_A$	$\widehat{S}_B$	$\widehat{S}_C$	$\widehat{S}_D$	$\widehat{S}_E$
$\widehat{S}_A$	–	49.3%	49.0%	38.9%	37.2%	$\widehat{S}_A$	–	19.1%	17.8%	4.4%	11.9%
$\widehat{S}_B$	49.3%	–	52.7%	43.5%	39.8%	$\widehat{S}_B$	19.1%	–	21.2%	7.7%	15.2%
$\widehat{S}_C$	49.0%	52.7%	–	41.6%	40.2%	$\widehat{S}_C$	17.8%	21.2%	–	6.3%	16.1%
$\widehat{S}_D$	38.9%	43.5%	41.6%	–	34.3%	$\widehat{S}_D$	4.4%	7.7%	6.3%	–	16.1%
$\widehat{S}_E$	37.2%	39.8%	40.2%	34.3%	–	$\widehat{S}_E$	11.9%	15.2%	16.1%	16.1%	–

Table 3. Superpixel group level average intersection rates of different scribbles on foreground and background.

foreground						background					
	$\widetilde{S}_A$	$\widetilde{S}_B$	$\widetilde{S}_C$	$\widetilde{S}_D$	$\widetilde{S}_E$		$\widetilde{S}_A$	$\widetilde{S}_B$	$\widetilde{S}_C$	$\widetilde{S}_D$	$\widetilde{S}_E$
$\widetilde{S}_A$	–	77.7%	78.3%	73.2%	77.6%	$\widetilde{S}_A$	–	66.9%	67.1%	58.2%	68.1%
$\widetilde{S}_B$	77.7%	–	81.5%	78.4%	79.8%	$\widetilde{S}_B$	66.9%	–	68.6%	61.7%	70.2%
$\widetilde{S}_C$	78.3%	81.5%	–	78.4%	78.8%	$\widetilde{S}_C$	67.1%	68.6%	–	60.6%	70.5%
$\widetilde{S}_D$	73.2%	78.4%	78.4%	–	75.3%	$\widetilde{S}_D$	58.2%	61.7%	60.6%	–	66.1%
$\widetilde{S}_E$	77.6%	79.8%	78.8%	75.3%	–	$\widetilde{S}_E$	68.1%	70.2%	70.5%	66.1%	–

decompose R, G, B channels into eight parts and the whole RGB color space is decomposed into $8 \times 8 \times 8$ subspaces. All the superpixels whose average color belong to the same subspace are considered as a superpixel group. Similarly, if one or more pixels in a superpixel group are labelled by a scribble, we consider the superpixel group to be labelled by the scribble.

Table 3 shows the intersection rates of the scribbles labelled by different users on superpixel group level. Here, $\widetilde{S}_n$ denotes the scribbles labelled by user n on all the images on superpixel group level, and intersection rates of different scribbles are calculated in a similar way to intersection rates on pixel level and superpixel level. We can find that the consistency of scribbles on superpixel group level keeps increasing, and the values of intersection rates are rather high on both foreground and background.

Based on the observation in Table 2 and 3, we conclude that the scribbles labelled by different users are consistent on superpixel group level but keep some varieties on superpixel level. The reason is that the key characteristics of foreground and background in each image are limited, and each key characteristic is represented by multiple superpixels especially on background. When a user labels an image, he/she usually tries to cover all the key characteristics of foreground and background to avoid further providing more interaction. Hence, most superpixel groups are labelled on both foreground and background. Nevertheless, the selection of superpixels to represent each key characteristic highly depends on personal habits. When a key characteristic is represented by multiple superpixels, the scribbles will appear obvious inconsistency.

Table 4. Content coverage rate by superpixel groups on foreground and background.

foreground						background					
	$\widetilde{S}_A$	$\widetilde{S}_B$	$\widetilde{S}_C$	$\widetilde{S}_D$	$\widetilde{S}_E$		$\widetilde{S}_A$	$\widetilde{S}_B$	$\widetilde{S}_C$	$\widetilde{S}_D$	$\widetilde{S}_E$
mean	74.9%	81.9%	83.6%	83.2%	76.5%	mean	64.9%	64.5%	65.4%	61.2%	74.1%
variance	0.03	0.03	0.02	0.02	0.03	variance	0.03	0.02	0.02	0.03	0.02
CV	0.22	0.17	0.17	0.18	0.22	CV	0.27	0.21	0.22	0.31	0.17

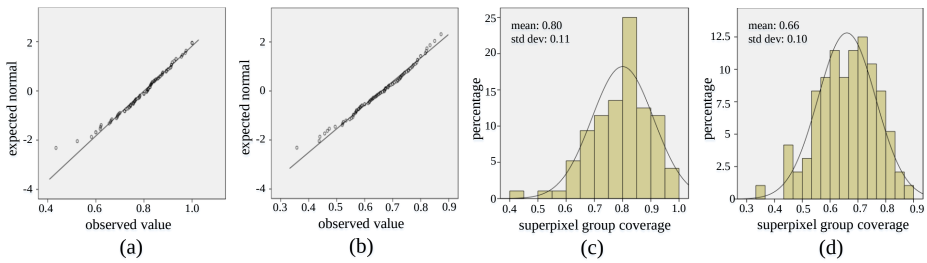


Fig. 3. Distribution of superpixel group coverage on foreground and background.

5.2 Distribution of superpixel group coverage

To effectively simulate the scribbles generated by different users, one important problem is how much image content should be covered by scribbles on superpixel group level, i.e., which percentage of superpixel groups should be labelled in scribble simulation. To answer this question, we calculate the percentages of superpixel groups labelled as foreground and background by different users. Table 4 shows the mean, variance and coefficient of variation (CV) of content coverage rate by superpixel groups on foreground and background, respectively. We can find that the content coverage rates of the scribbles labelled by different users only have small differences from similar mean values, and the content coverage rates of the scribbles labelled by the same user are stable from low values of variance and CV. Hence, in simulating the scribbles labelled by one user, the content coverage rate can be randomly selected in a small range but it should be keep consistent in scribble simulation on all images.

To describe the distribution of superpixel group coverage, we test its normality with normal Q-Q plot. Fig. 3 (a) and (b) show the normal Q-Q plots of superpixel group coverage on foreground and background, respectively. It shows that data points in the plots are both close to the diagonals, which indicates that the distribution of superpixel group coverage on foreground and background are both normal distribution. We analyze the parameters of these two normal distributions in Fig. 3 (c) and (d), and use them to describe the distribution of superpixel group coverage on foreground and background.

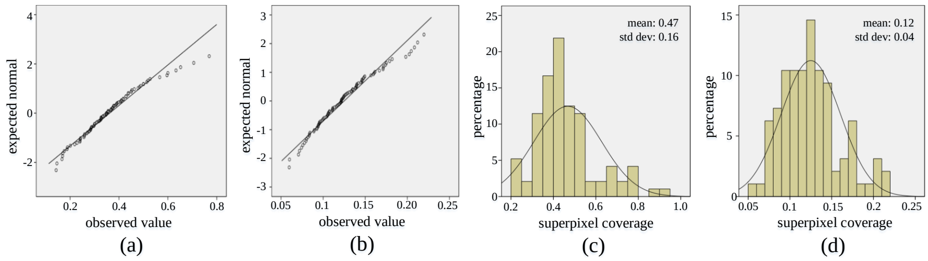


Fig. 4. Distribution of superpixel coverage on foreground and background.

5.3 Distribution of superpixel coverage

Another problem in scribble simulation is how to describe the distribution of superpixel coverage since it has obvious inconsistency among the scribbles labelled by different users.

We analyze the distribution of superpixel coverage in a similar way to superpixel group coverage. Fig. 4 (a) and (b) show the normal Q-Q plots of superpixel coverage on foreground and background, respectively. It shows that the distribution of superpixel coverage on foreground and background are also both normal distribution. And Fig. 4 (c) and (d) show the parameters of these two normal distributions, which are used to describe the distribution of superpixel coverage on foreground and background.

5.4 Effect of connection in scribble

The third problem in scribble simulation is how to generate the smooth curves to represent the scribbles which should look natural and cover the prescribed percentage of superpixel groups. It is a difficult and complex problem though it has been researched for decades [20–22]. To simplify the problem, we analyze the effective elements in scribble for interactive image segmentation.

To each manually labelled scribble, we randomly select one pixel from each superpixel covered by the scribble. In this way, we obtain a pixel set as the representative of each scribble. Then, we generate the segmentation results using the scribbles and their corresponding pixel sets as the inputs respectively, and compare the segmentation results by the criteria of precision and recall.

Fig. 5 shows the comparison of the segmentation results generated from the inputs of manually labelled scribbles (blue bins) and their corresponding pixel sets without connections (orange bins). It shows that the segmentation results generated by these two types of inputs are very similar in performance. It means that the effect of connection in scribbles is weak to interactive image segmentation. Hence, we can use arbitrary connections between pixels in scribble simulation when keeping the stability of superpixel group coverage, or even completely ignore the connections. In our experiments, we directly use pixel sets without connections to simulate scribbles to simplify processing procedure and reduce computational cost.

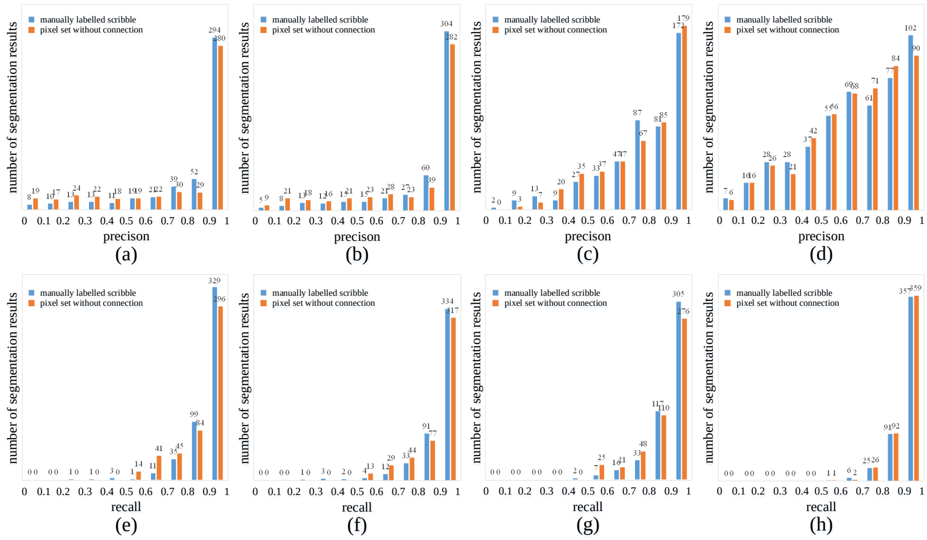


Fig. 5. Comparison of the segmentation results generated by manually labelled scribbles and their corresponding pixel sets. (a)-(d) Precision comparison using SC, GSC, RW, and GSP algorithms, respectively. (e)-(h) Recall comparison using SC, GSC, RW, and GSP algorithms, respectively.

5.5 Scribble simulation

Based on the above analysis, we simulate scribbles on foreground and background separately according to the segmentation ground truth of each image. We first generate the superpixels and superpixel groups according to the corresponding segmentation ground truth. Then, we determine the values of superpixel group coverage and superpixel coverage according to the distribution in Fig. 3 and 4. Thereafter, we randomly select superpixels groups until reaching the determined superpixel group coverage value. Finally, we randomly selected superpixels within the selected superpixel groups, in which at least one superpixel is selected within each selected superpixel group and the total number of all the superpixels selected within in all groups is equal to the determined value of superpixel coverage.

6 Evaluation of Interactive Segmentation Algorithms

To validate the performance of the proposed scribble simulation approach, we compare the segmentation results generated by manually labelled scribbles and automatically simulated scribbles on the 96 images in our data set. To illustrate the effectiveness of our approach, we also use two other scribble simulation approaches in comparison, randomly selecting one superpixel in each selected superpixel group (SPG) and randomly selecting superpixels without grouping

(SP). For each image is labelled by five users, we simulate five scribbles for each image with any simulation approach.

We evaluate the scribbles on all the four interactive image segmentation algorithms, including graph cuts (GC) [8], geodesic star convexity (GSC) [12], random walker (RW) [11], and geodesic shortest path (GSP) [14]. Fig. 6 shows the evaluation results on the criteria of precision and recall. It shows that the evaluation results based on the scribbles generated by our approach is similar to the ones based on manually labelled scribbles. And it outperforms the other two simulation approaches in avoiding serious deviation in precision and recall evaluation (Fig. 6 (b)(c)(g)).

7 Conclusion

In this paper, we propose an automatic scribble simulation approach for interactive segmentation algorithm evaluation. Based on the analysis of the scribble variety, we describe the consistency and inconsistency of scribbles with normal distribution on superpixel level and superpixel group level, and simulate scribbles on foreground and background separately by randomly selecting superpixel groups and superpixels with the previously determined coverage values. The experimental results evaluated by four existing interactive image segmentation algorithms on 96 images show that the scribbles simulated by the proposed approach can obtain similar evaluation results to manually labelled scribbles and avoid serious deviation in both precision and recall.

Acknowledgments. This work is supported by the National Science Foundation of China (61321491, 61202320), Research Project of Excellent State Key Laboratory (61223003), National Undergraduate Innovation Project (G1410284075) and Collaborative Innovation Center of Novel Software Technology and Industrialization.

References

1. Bing-Kun Bao, Guangcan Liu, Richang Hong, Shuicheng Yan, and Changsheng Xu. General subspace learning with corrupted training data via graph embedding. IEEE TIP. 22(11):4380–4393 (2013)
2. Xiangyang Xu, Wenjing Geng, Ran Ju, Yang Yang, Tongwei Ren, Gangshan Wu. OBSIR: Object-based stereo image retrieval. IEEE ICME. 1-6 (2014)
3. Jitao Sang, Changsheng Xu, Jing Liu. User-aware image tag refinement via ternary semantic analysis. IEEE TMM. 14(3-2):883-895 (2012)
4. Teng Li, Huan Chang, Meng Wang, Bingbing Ni, Richang Hong, and Shuicheng Yan. Crowded scene analysis: A survey. IEEE TCSVT. 25(3), 367–386 (2015)
5. Tongwei Ren, Zhongyan Qiu, Yan Liu, Tong Yu, and Jia Bei. Soft-assigned bag of features for object tracking. MMSJ. 21(2):189-205 (2015)
6. Jitao Sang, Changsheng Xu, and Jing Liu. User-Aware image tag refinement via ternary semantic analysis. IEEE TMM. 14(3-2):883–895 (2012)

7. Bing-Kun Bao, Guangcan Liu, Changsheng Xu, and Shuicheng Yan. Inductive robust principal component analysis. IEEE TIP. 21(8):3794–3800 (2012)
8. Yuri Y Boykov, and Marie-Pierre Jolly. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. IEEE ICCV. 105–112 (2001)
9. Ran Ju, Tongwei Ren, and Gangshan Wu. StereoSnakes: Contour based consistent object extraction for stereo images. IEEE ICCV (2015)
10. Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. Grabcut: Interactive foreground extraction using iterated graph cuts. ACM TOG. 23(3):309–314 (2004)
11. Leo Grady. Random walks for image segmentation. IEEE TPAMI. 28(11):1768-1783 (2006)
12. Varun Gulshan, Carsten Rother, Antonio Criminisi, Andrew Blake, and Andrew Zisserman. Geodesic star convexity for interactive image segmentation. IEEE CVPR. 3129–3136 (2010)
13. Ling Ge, Ran Ju, Tongwei Ren, and Gangshan Wu. Interactive RGB-D image segmentation using hierarchical graph cut and geodesic distance. PCM. 114–124 (2015)
14. Xue Bai and Guillermo Sapiro. A geodesic framework for fast interactive image and video segmentation and matting. IEEE ICCV. 1–8 (2007)
15. Kevin McGuinness and Noel E O'Connor. A comparative evaluation of interactive segmentation algorithms. PR. 43(2):434–444 (2010)
16. David Martin, Charless Fowlkes, Doron Tal, and Jitendra Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. IEEE ICCV. 416–423 (2001)
17. Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, and Eero P. Simoncelli. Image quality assessment: from error visibility to structural similarity. IEEE TIP. 13(4):600-612 (2004)
18. Tongwei Ren, and Gangshan Wu. Automatic image retargeting evaluation based on user perception. IEEE ICIP. 1569–1572 (2010)
19. Tongwei Ren, Yan Liu, and Gangshan Wu. Full-reference quality assessment for video summary. IEEE ICDM Workshops. 874–883 (2008)
20. Yu Fu, Jian Cheng, Zhenglong Li, and Hanqing Lu. Saliency cuts: An automatic approach to object segmentation. ICPR. 1–4 (2008)
21. Kevin McGuinness and Noel O'Connor. Toward automated evaluation of interactive segmentation. CVIU. 115(6):868–884 (2011)
22. Pushmeet Kohli, Hannes Nickisch, Carsten Rother, and Christoph Rhemann. User-centric learning and evaluation of interactive segmentation systems. IJCV. 100(3):261–274 (2012).
23. Emmanouil Moschidis and Jim Graham. A systematic performance evaluation of interactive image segmentation methods based on simulated user interaction. IEEE ISBI. 928–931 (2010)
24. Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Susstrunk. Slic superpixels compared to state-of-the-art superpixel methods. IEEE TPAMI. 34(11):2274–2282 (2012)
25. Andrea Vedaldi and Brian Fulkerson. VLFeat: An open and portable library of computer vision algorithms. ACM MM. 1469–1472 (2010)

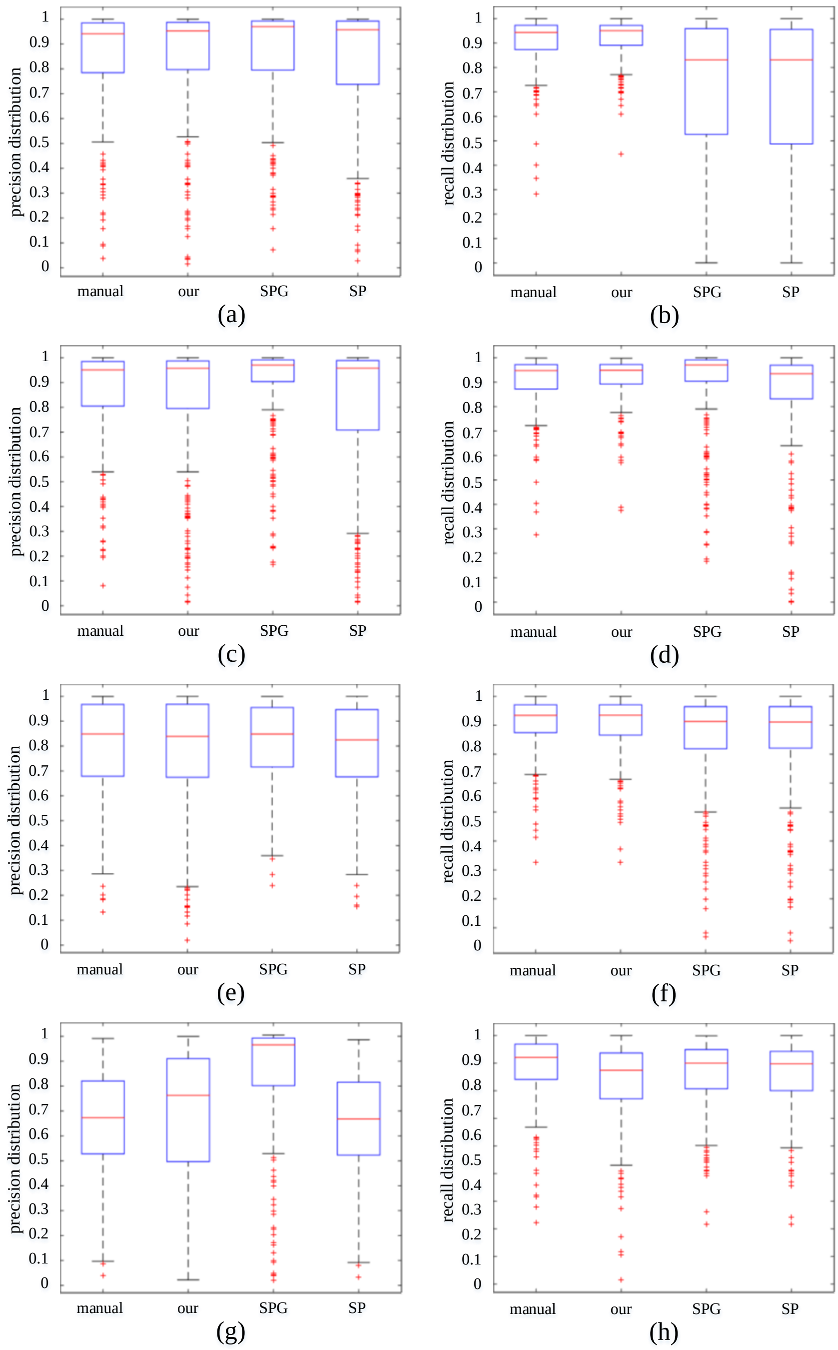


Fig. 6. Comparison of the evaluation results. (a)(c)(e)(g) Precision evaluation the segmentation results generated by manually labelled scribbles and the scribbles simulated by our approach, SPG, SP, respectively. (b)(d)(f)(h) Recall evaluation the segmentation results generated by manually labelled scribbles and the scribbles simulated by our approach, SPG, SP, respectively.