International Journal on Advanced Science Engineering Information Technology

Vol.10 (2020) No. 1
ISSN: 2088-5334

A Comparative Study of Interactive Segmentation with Different Number of Strokes on Complex Images

Kok Luong Goh[#], Giap Weng Ng[#], Muzaffar Hamzah[#], Soo See Chai[*]

[#] *Faculty of Computing & Informatics, University Malaysia Sabah, Kota Kinabalu, 88400, Malaysia*
E-mail: luonggoh@yahoo.com; nggiapweng@ums.edu.my; muzaffar@ums.edu.my

[*] *Faculty of Computer Science & Information Technology, University Malaysia Sarawak, Kota Samarahan, 94300, Malaysia*
E-mail: sschai@unimas.my

***Abstract*— Interactive image segmentation is the way to extract an object of interest with the guidance of the user. The guidance from the user is an iterative process until the required object of interest had been segmented. Therefore, the input from the user as well as the understanding of the algorithms based on the user input has an essential role in the success of interactive segmentation. The most common user input type in interactive segmentation is using strokes. The different number of strokes are utilized in each different interactive segmentation algorithms. There was no evaluation of the effects on the number of strokes on this interactive segmentation. Therefore, this paper intends to fill this shortcoming. In this study, the input strokes had been categorized into single, double, and multiple strokes. The use of the same number of strokes on the object of interest and background on three interactive segmentation algorithms: i) Nonparametric Higher-order Learning (NHL), ii) Maximal Similarity-based Region Merging (MSRM) and iii) Graph-Based Manifold Ranking (GBMR) are evaluated, focusing on the complex images from Berkeley image dataset. This dataset contains a total of 12,000 test color images and ground truth images. Two types of complex images had been selected for the experiment: image with a background color like the object of interest, and image with the object of interest overlapped with other similar objects. This can be concluded that, generally, more strokes used as input could improve image segmentation accuracy.**

***Keywords*— image segmentation; interactive segmentation; user input; strokes; complex image.**

I. INTRODUCTION

Image segmentation is a process to help human to extract object of interest from an image. Image segmentation can be categorized into manual, semi-automated and fully automated. There is no involvement of users in the automated segmentation. On the other hand, semi-automated, also known as interactive segmentation, requires minimal user intervention during the segmentation process. Image segmentation aims to automate the whole image segmentation process fully. However, automated segmentation still facing huge obstacles in producing satisfactory results due to the complexity of the images. Therefore, semi-automated or interactive image segmentation are preferred ways to achieve better results in the image segmentation.

Interactive image segmentation had been used in various applications. For example, tools had been developed for medical volume images (SmartPaint [1] and MRI for orthopedic surgery [2]). Besides that, a segmentation tool had been developed for lithological boundary detection in remote sensing areas [3]. Furthermore, interactive segmentation has been playing an essential role in agriculture by helping farmers for crop disease detection [4].

As explained in the previous paragraph, the user will guide the segmentation system to extract the object of interest in interactive segmentation. The general process of interactive segmentation is summarized as below:

- Step 1: The user will provide information on the background and object of interest.
- Step 2: The segmentation system will produce the segmentation result based on the input from the user.
- Step 3: The user will evaluate the result and the whole process will stop if the user is satisfied with the result. Otherwise, the user will continue to provide additional information (background and object of interest) until the system has produced a satisfactory segmented result.

A good interactive segmentation system will be able to produce a satisfactory result with minimal input information from the users. In order to achieve this, the system should be designed to understand the intended meaning of the user input.

There are different input types used in the interactive segmentation to provide information on the background and object of interest. Below are the examples of these input types:

- Stroke(s) [5-8]: the user is required to place stroke(s) on the object of interest and background on the image
- Bounding box [9-11]: the user is required to put the bounding box on the object of interest in the image.
- Seed point [12, 13]: the user is required to put the seed points on the background and object of interest in the image.

Besides above-mentioned input types, [14] had employed placing seed points on the contour of the object of interest. In addition, [15, 16] had applied a combination of strokes and seed points in the segmentation process. Fig. 1 shows these input types in interactive segmentation.

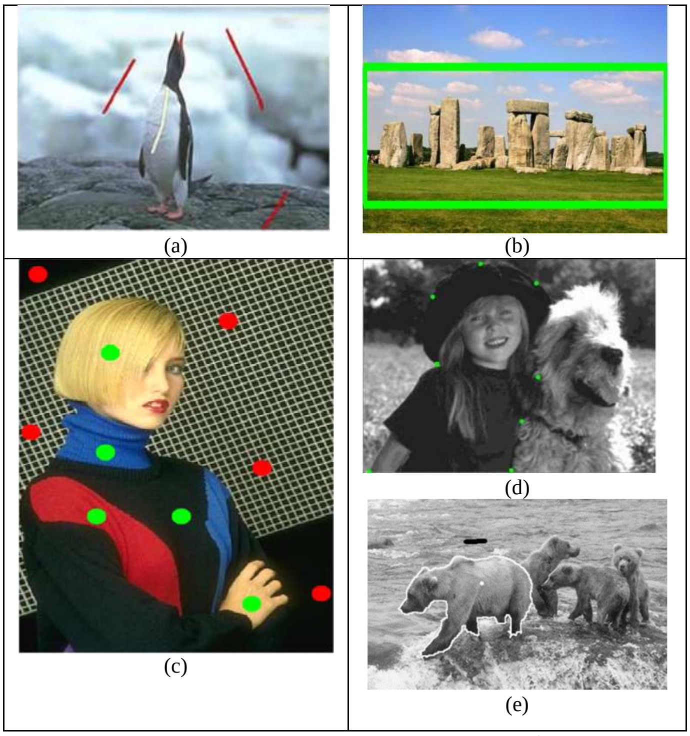


Fig. 1. Interactive segmentation user input types: (a) foreground and background strokes, (b): bounding boxes for object of interest. (c) Seed points for the object of interest and background of the image. (d): placing seed points on the contour of the object of interest. (e): combination of seed points and strokes for the object of interest and background of the image.

The use of stroke(s) is the most common input type in interactive segmentation. However, it was noticed that, different numbers of strokes were applied in different algorithms and the effects of this are not being addressed. In other words, there is no study on evaluation on the number of strokes on the different interactive algorithms to verify the effects of this in the accuracy obtained. In our previous work [17], it was reported that the location, number of inputs and length of the inputs would affect the retrieval accuracy on complex image while remain consistent for simple image. This paper intends to extend the previous work by comparing the effects on the different number of strokes on three interactive segmentation algorithms: Nonparametric Higher-order Learning for interactive segmentation (NHL) [17], interactive image segmentation by Maximal Similarity-based Region Merging (MSRM) [18], and robust interactive image segmentation via Graph-Based Manifold Ranking (GBMR) [19]. Below are the brief descriptions of the three algorithms:

A. *Nonparametric Higher-order Learning for interactive segmentation (NHL) [17]*

This is a generative interactive model which is used to estimate the likelihood of the pixel for each label. A new high-order cost function of pixel likelihoods used to enforce the labeling consistency was designed by using the mean shift unsupervised learning algorithm. Multiple over-segmentation is needed to be applied to this algorithm in order to obtain a good segmentation result. Fig. 2 shows the strokes used to indicate the object of interest and background, over segmented image, and result generated by the algorithm.

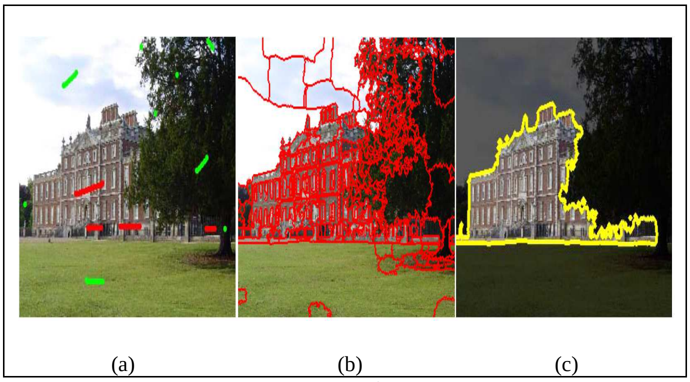


Fig. 2. Segmentation process obtained from algorithm [17]: (a) strokes input by users. (b) oversegmented image. (c) the segmentation results

B. *Interactive image segmentation by Maximal Similarity-based Region Merging (MSRM) [18]*

Maximal-similarity algorithm was based on the region merging method. In this algorithm, the image will initially over segmented by mean shift segmentation. The algorithm will next extract the object contour by labeling all the non-marker regions as either background or object of interest. Fig. 3 shows the input information on the oversegmented image and the result obtained by using this algorithm.

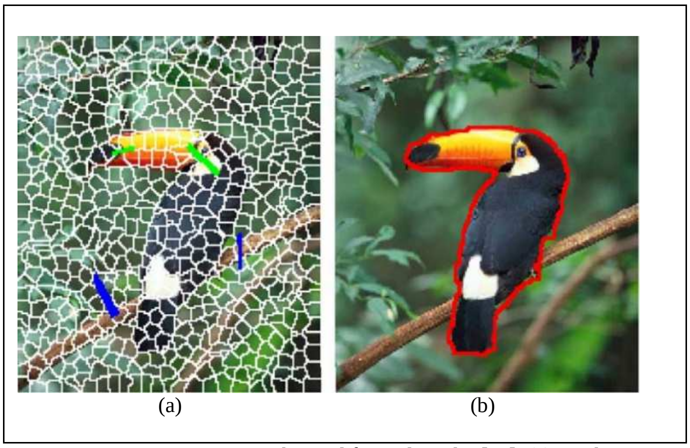


Fig. 3. Segmentation process obtained from algorithm[18]: (a) strokes input by users on the over segmented image. (b): the segmentation result

C. *Robust interactive image segmentation via Graph-based Manifold Ranking (GBMR)* [19]

This algorithm is based on the approximately of *k*-regular sparse graph which forms the affinity graph matrix using driven labels and locally adaptive kernel parameters. User input information is integrated into oversegmented images and the output of the segmentation is generated by using integration of the output from the learning of background and foreground labels. Fig. 4 shows the background and foreground labels on the oversegmented image and the output obtained.

(a) (b)

Fig. 4. Segmentation process obtained from algorithm [17]: (a) foreground and background labels on the over segmented image. (b): the segmentation result

II. MATERIALS AND METHOD

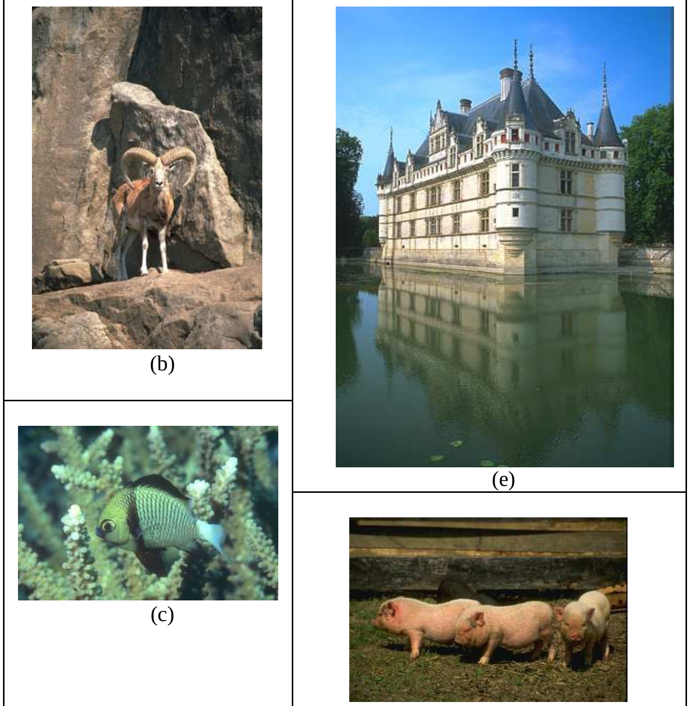
In the previous section, the three algorithms use strokes to label the object of interest and background in the image with the different number of strokes used for the background and object of interest. It was also noticed that, from the review of these three interactive segmentation algorithms, there were little or no evaluation of the algorithms on the complex images. Based on these observations, this paper aims to evaluate these three algorithms: 1) on the complex images by using stroke input, and 2) on the different numbers of input strokes. Complex images are defined as:

- image with a background color like the object of interest (image (a) to (e) in Fig. 5), and
- image with the object of interest overlapped with other similar objects (image (f) in Fig. 5).

Some of the complex images selected from the Berkeley Segmentation Dataset [20] for the testing in this paper are shown in Fig. 5. The ground truth of these selected images is included in Fig. 6.

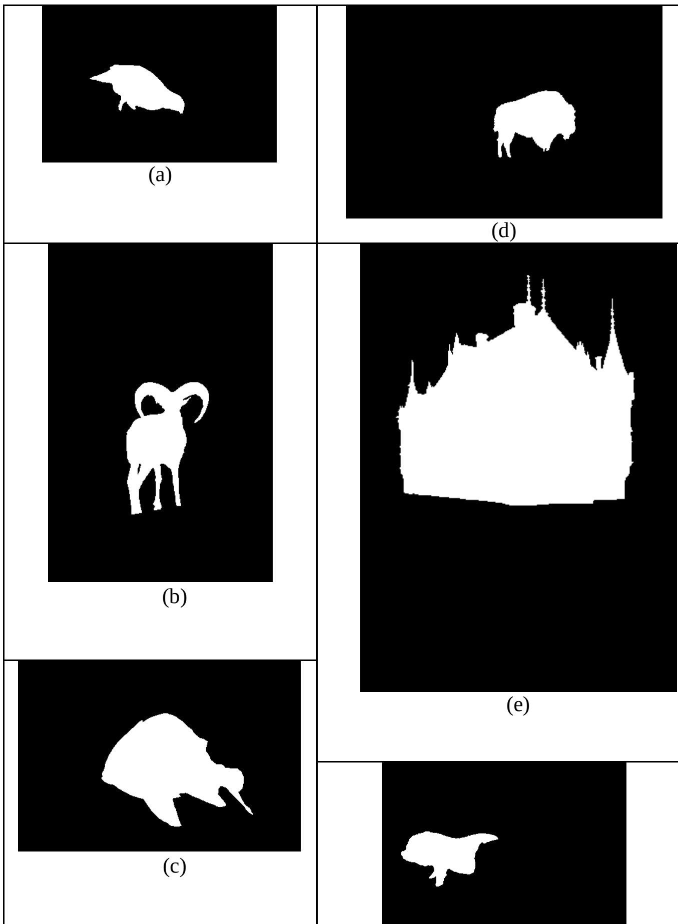
Quantitative evaluation on the segmentation results will be done with three evaluation parameters: Variation of Information (VI), Global Consistency Error (GCE) and Jaccard Index (JI) [21]. VI provides the distance information between the segmentation result and the ground truth. GCE measures of dissimilarities between the ground truth and the segmented image. Lastly, JI measures the percentage of overlap between the ground truth and the segmentation result.

(a) (d)



(b) (e) (c) (f)

Fig. 5. (a)-(e): image with a background color like the object of interest, and (f): image with the object of interest overlapped with other similar objects.



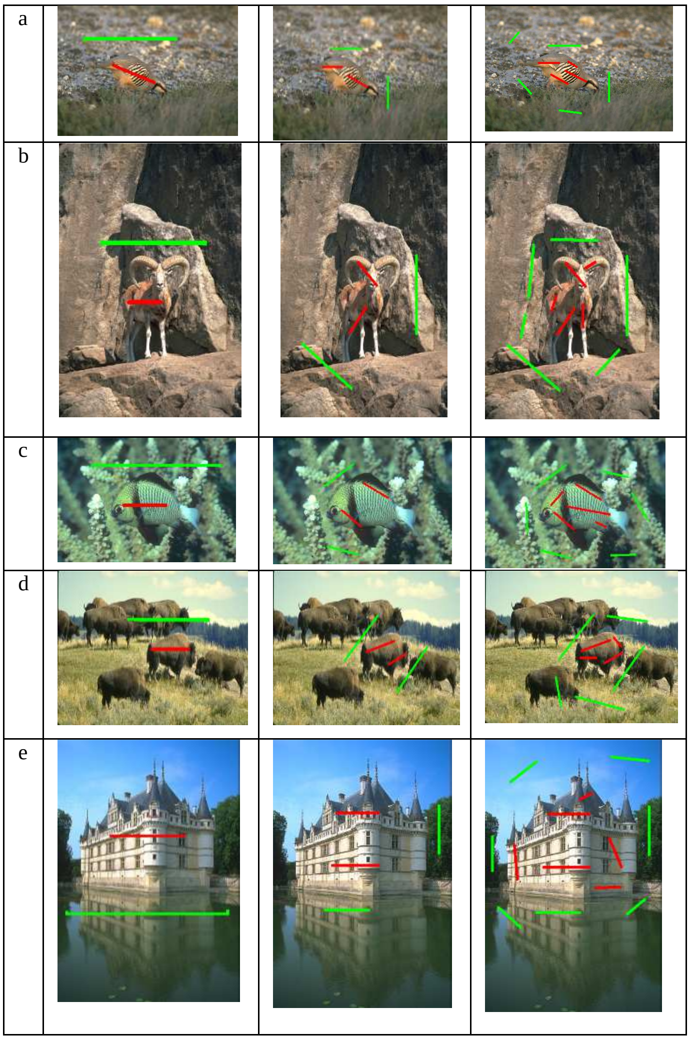
(a) (d) (b) (e) (c) (f)

Fig. 6. Ground truth for the test images

For the three algorithms, strokes are being applied to represent the background information and object of interest. In order to assess the effect of the number of strokes on these three algorithms, we had divided input strokes into three categories (Fig. 7) for:

- Single strokes: one stroke on the background and one stroke on the object of interest,
- Double strokes: two strokes on the background and two strokes on the object of interest, and
- Multiple strokes: more than two strokes on the background and more than two strokes on the object of interest

Besides the number of strokes used for each algorithm, the original over-segmentation technique used in each of the algorithms remains. Six (6) complex images of two categories: image with a background color like the object of interest (a. to e.) and image with the object of interest overlapped with other similar objects (f) as shown in Fig. 5 will be used. For these six complex images, the different number of strokes, i.e., single strokes, double strokes, and multiple strokes as will be tested. The location and length for each of these different numbers of strokes are the same for all the three algorithms, as shown in Fig. 7. The values of the evaluation parameters: GCE, VI, and JI, will be calculated for each image.



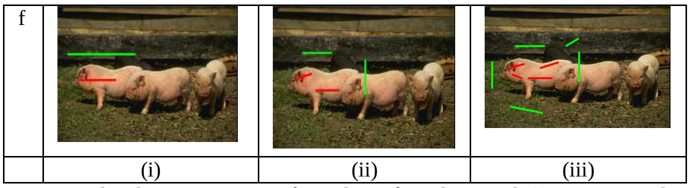


Fig. 7. The three categories of number of strokes used (i): images with single stroke, (ii): image with double strokes, and (iii): image with multiple strokes for image a. to f.

III. Results and Discussion

For better evaluation of the results obtained, the analysis is based on the different number of strokes used.

A. *Single Stroke*

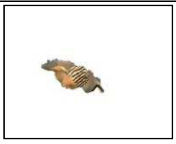
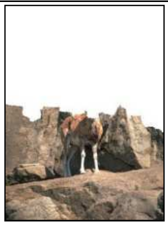
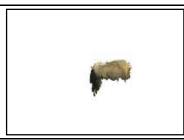
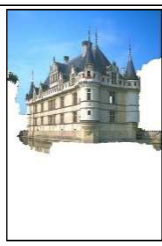
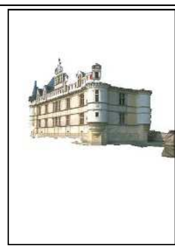
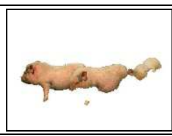
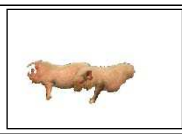
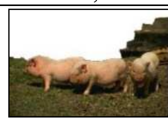
NHL	MRSM	GBMR
		
GCE=0.03, VI=0.97, JI=0.74	GCE=0.12, VI=0.55, JI=0.08	GCE=0.12, VI=0.50, JI=0.10
		
GCE=0.03, VI=0.92, JI=0.39	GCE=0.09, VI=0.87, JI=0.28	GCE=0.13, VI=0.50, JI=0.10
		
GCE=0.26, VI=0.52, JI=0.21	GCE=0.08, VI=0.89, JI=0.63	GCE=0.25, VI=0.51, JI=0.25
		
GCE=0.03, VI=0.96, JI=0.65	GCE=0.09, VI=0.52, JI=0.08	GCE=0.09, VI=0.52, JI=0.08
		
GCE=0.26, VI=0.60, JI=0.48	GCE=0.26, VI=0.61, JI=0.48	GCE=0.30, VI=0.56, JI=0.41
		
GCE=0.08, VI=0.84, JI=0.42	GCE=0.08, VI=0.86, JI=0.42	GCE=0.12, VI=0.51, JI=0.11

Fig. 8. Segmentation result generated from (a) NHL[17], (b) MSRM[18] and (c) GBMR by using single stroke for input each object and background.

B. Double strokes

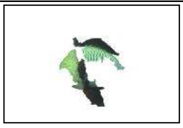
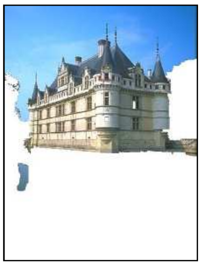
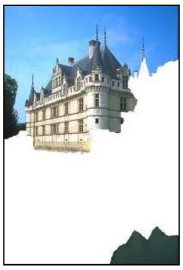
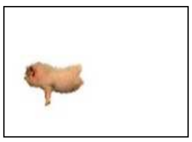
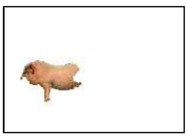
NHL	MRSM	GBMR
		
GCE=0.03 VI=0.96 JI=0.72	GCE=0.03 VI=0.96 JI=0.72	GCE=0.12 VI=0.51 JI=0.13
		
GCE=0.05 VI=0.93 JI=0.59	GCE=0.09 VI=0.87 JI=0.40	GCE=0.12 VI=0.50 JI=0.12
		
GCE=0.04 VI=0.76 JI=0.13	GCE=0.09 VI=0.85 JI=0.49	GCE=0.25 VI=0.52 JI=0.24
		
GCE=0.07 VI=0.81 JI=0.32	GCE=0.06 VI=0.86 JI=0.40	GCE=0.09 VI=0.51 JI=0.10
		
GCE=0.26 VI=0.61 JI=0.48	GCE=0.27 VI=0.61 JI=0.47	GCE=0.35 VI=0.53 JI=0.33
		
GCE=0.03 VI=0.97 JI=0.79	GCE=0.02 VI=0.97 JI=0.81	GCE=0.12 VI=0.50 JI=0.11

Fig. 9. Segmentation result generated from (a) NHL[17], (b) MSRM[18] and (c) GBMR by using double strokes for each object and background.

C. Multiple Strokes

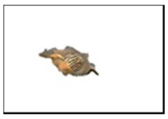
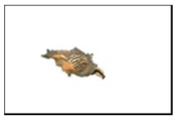
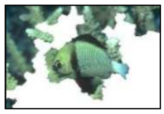
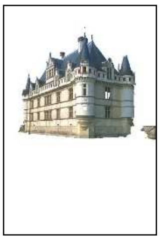
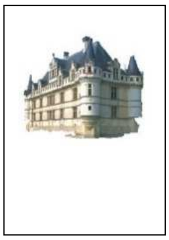
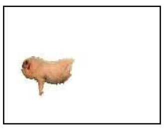
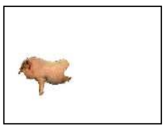
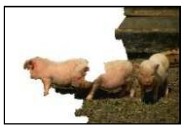
NHL	MRSM	GBMR
		
GCE=0.02 VI=0.98 JI=0.83	GCE=0.02 VI=0.98 JI=0.83	GCE=0.11 VI=0.54 JI=0.15
		
GCE=0.04 VI=0.95 JI=0.71	GCE=0.03 VI=0.97 JI=0.77	GCE=0.12 VI=0.52 JI=0.14
		
GCE=0.08 VI=0.84 JI=0.44	GCE=0.05 VI=0.94 JI=0.82	GCE=0.23 VI=0.53 JI=0.29
		
GCE=0.04 VI=0.93 JI=0.60	GCE=0.02 VI=0.98 JI=0.81	GCE=0.09 VI=0.54 JI=0.12
		
GCE=0.03 VI=0.97 JI=0.95	GCE=0.03 VI=0.97 JI=0.94	GCE=0.29 VI=0.57 JI=0.44
		
GCE=0.03 VI=0.97 JI=0.78	GCE=0.02 VI=0.97 JI=0.81	GCE=0.12 VI=0.50 JI=0.11

Fig. 10. Segmentation result generated from (a) NHL [17], (b) MSRM [18] and (c) GBMR by using multiple strokes for each object and background.

Fig. 8 to 10 show the results generated from the three interactive segmentation algorithms by using three different numbers of strokes ranging from single, double and multiples. The findings of the testing are summarized as below:

- Regardless of the number of strokes, MSRM and NHL can achieve averagely low GCE (0.08) and high VI (0.85) as compared to GBMR (GCE=0.17 and VI=0.52) for the complex images. In between MSRM and NHL, MSRM performed slightly better than NHL.
- Based on the JI measurement, MSRM performed slightly weak in image b. NHL, on the other hand, was weak in images b, c and d.
- In terms of number of the input stroke category, multiple strokes performed better than double and single stroke regardless of the interactive segmentation algorithms.
- The double stroke information is enough to achieve a result as good as multiple strokes in image f.
- NHL generated a good result by using single as compared to double and multiple strokes in image d.
- The results produced by using a single stroke are better than double strokes used in the image c by using MSRM algorithm.
- NHL produced a better result with an average GCE=0.12, VI=0.8, and JI=0.48 as comparing to MSRM and GBMR by using single stroke input.
- MSRM performed slightly better by using double strokes input with average GCE=0.09, VI=0.85, and JI=0.55 as compared with NHL and GBMR.
- MSRM outperformed the other two algorithms by using multiple strokes input.

For images with a background color like the object of interest (a-e), multiple strokes are required in order to achieve a good segmentation result. However, for image (f) whereby the object of interest overlapped with other similar objects, double strokes had shown a good segmentation result. This could be due to:

- The interactive segmentation algorithms could not differentiate between the object of interest and background when single and double strokes were used in images with a background color like the object of interest. Therefore, more strokes were required in order to obtain a good segmentation result.
- By using multiple strokes, additional information on the color and location of the object of interest would be better obtained when segmenting the images with a background color like the object of interest.
- For the image with the object of interest overlapped with other similar objects, double strokes were found to segment the object of interest successfully. The use of multiple strokes did not show to increase accuracy more. This could be since there was a huge contrast on the background and the object of interest in the overlapping object image and the use of double strokes managed to include the location of the object of interest precisely.

In terms of the performance of the algorithm, MSRM had performed better than NHL and GBMR. On the other hand, the findings of this experiment had shown that more input is required by the interactive algorithms in order to achieve a better result. This is against the aim of the interactive segmentation, i.e., a good interactive segmentation system will be able to produce a satisfactory result with minimal input information from the users. Therefore, further research is required to improve the interactive segmentation algorithm with minimal input. Besides that, the finding of a suitable input type could be another way to improve the result of segmentation.

IV. Conclusion

This paper evaluates the three interactive segmentation algorithms: NHL, MSRM, and GBMR with three categories of a few input strokes on complex images. Variation of Information (VI), Global Consistency Error (GCE) and Jaccard Index (JI) had been selected to evaluate the output generated from these three interactive segmentation algorithms. MSRM had produced a better segmentation result as compared to NHL and GBMR in the complex images. In terms of the number of strokes, the results produced by using multiple strokes outperformed double and single strokes used due to the additional information fed to the interactive segmentation algorithms, especially in images with a background color like the object of interest.

References

[1] Malmberg, F., et al., *SmartPaint: a tool for interactive segmentation of medical volume images.* Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 2017. **5**(1): p. 36-44.

[2] Ozdemir, F., et al., *Interactive segmentation in MRI for orthopedic surgery planning: bone tissue.* International Journal of Computer Assisted Radiology and Surgery, 2017. **12**(6): p. 1031-1039.

[3] Vasuki, Y., et al., *An interactive image segmentation method for lithological boundary detection: A rapid mapping tool for geologists.* Computers & Geosciences, 2017. **100**: p. 27-40.

[4] Ma, J., et al., *A segmentation method for greenhouse vegetable foliar disease spots images using color information and region growing.* Computers and Electronics in Agriculture, 2017. **142**: p. 110-117.

[5] Zhou, C., et al., *An efficient two-stage region merging method for interactive image segmentation.* Computers & Electrical Engineering, 2016. **54**: p. 220-229.

[6] Madani, E.M. and M. Torki. *Interactive Image Segmentation Using Multimodal Regularized Kernel Embedding*. in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*. 2018.

[7] Jian, M. and C. Jung, *Interactive Image Segmentation Using Adaptive Constraint Propagation.* IEEE Transactions on Image Processing, 2016. **25**(3): p. 1301-1311.

[8] Park, S., H.S. Lee, and J. Kim. *Seed growing for interactive image segmentation with geodesic voting*. in *2016 IEEE International Conference on Image Processing (ICIP)*. 2016.

[9] Yu, H., et al. *Loosecut: Interactive image segmentation with loosely bounded boxes*. in *2017 IEEE International Conference on Image Processing (ICIP)*. 2017.

[10] He, K., et al., *Interactive Image Segmentation on Multiscale Appearances.* IEEE Access, 2018. **6**: p. 67732-67741.

[11] Li, X., K. Liu, and Y. Dong, *Superpixel-Based Foreground Extraction with Fast Adaptive Trimaps.* IEEE Transactions on Cybernetics, 2018. **48**(9): p. 2609-2619.

[12] Feng, J., et al. *Interactive Segmentation on RGBD Images via Cue Selection*. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016.

[13] Li, Z., Q. Chen, and V. Koltun. *Interactive Image Segmentation with Latent Diversity*. in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018.

[14] Zadicario, E., et al. *Boundary snapping for robust image cutouts*. in *2008 IEEE Conference on Computer Vision and Pattern Recognition.* 2008.

[15] de Miranda, P.A.V., A.X. Falcão, and J.K. Udupa, *Synergistic arc-weight estimation for interactive image segmentation using graphs.* Computer Vision and Image Understanding, 2010. **114**(1): p. 85-99.

[16] Luo, L., et al., *Interactive image segmentation based on samples reconstruction and FLDA.* Journal of Visual Communication and Image Representation, 2017. **43**: p. 138-151.

[17] Kim, T.H., K.M. Lee, and S.U. Lee. *Nonparametric higher-order learning for interactive segmentation*. in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2010.

[18] Ning, J., et al., *Interactive image segmentation by maximal similarity-based region merging.* Pattern Recogn., 2010. **43**(2): p. 445-456.

[19] Li, H., W. Wu, and E. Wu, *Robust interactive image segmentation via graph-based manifold ranking.* Computational Visual Media, 2015. **1**(3): p. 183-195.

[20] Martin, D., et al. *A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics*. in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*. 2001.

[21] Sharma, M.a.V.C., *Objective evaluation parameters of image segmentation algorithms.* International Journal of Engineering and Advanced Technology (IJEAT), 2012. **2**(2): p. 2249-8958.