Interactive Segmentation as Supervised Classification with Superpixels

Han S. Lee, Jiwhan Kim, Sun Jeong Park, and Junmo Kim
Department of Electrical Engineering
Korea Advanced Institute of Science and Technology
hansanglee@kaist.ac.kr

1. Introduction

Image segmentation is a task of extracting desired objects from an image. Though automated image segmentation have been widely researched for decades, the results of automated segmentation is not yet satisfactory enough. One reason is that the "desired objects" in image segmentation are quite subjective so that there could be different objects desired for the same image depending on the purpose of task. As an alternative, an interactive segmentation takes ths human inputs called "seeds" to capture the desired objects by specifying the object and background from them. As a semi-automatic segmentation technique, interactive segmentation has also been actively researched in certain applications including medical image segmentation, which especially requires human expert's knowledge and interaction to complete the task.

Since the procedure of seeding costs time and human resources, the goal of interactive segmentation is to extract the accurate objects with the minimal human inputs. To achieve this, numerous researches have been studied in recent decades, including graph-based approaches such as interactive graph cuts [2] and Random Walks [3]. Besides the graph-based models, the learning-based models for interactive segmentation have also been introduced recently. The learning-based models view the image as a set of data i.e. pixels, and the interactive segmentation as a supervised classification of data according to the training set of seed pixels. Recently, number of these learning-based approaches showed promising results.

In this paper, we suggest the learning-based interactive segmentation method with supervised classification of superpixels. Viewing the interactive segmentation as a supervised classification problem, we construct the supervised classification framework with superpixels for interactive segmentation. Additionally, the label refinement in pixel-wise level is performed to refine the boundary. In experiments, our method showed not only promising but also competitive performances compared to the state-of-the-art interactive segmentation models.

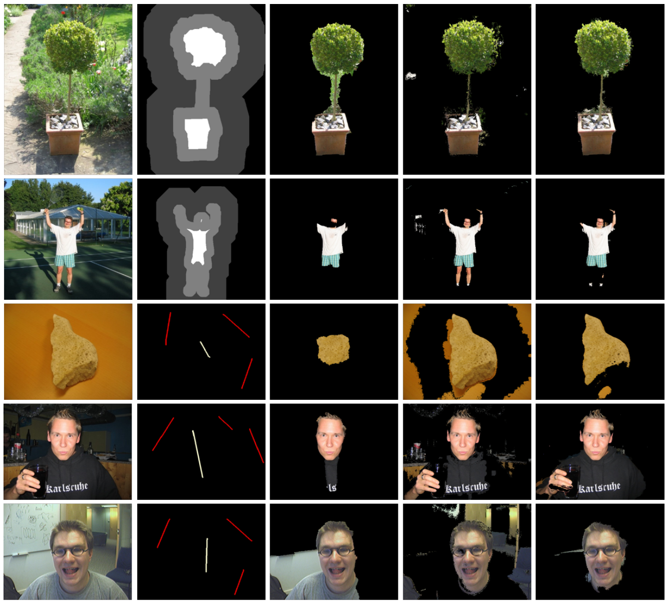
2. Proposed Method

The basic idea of our method is to interpret the interactive segmentation task as a supervised classification problem. In the learning-based models, we view each pixel, or superpixel in our case, and corresponding feature vector as a data to be classified into object or background. Specifically, user-labeled seed pixels are interpreted as labeled data which are to be used as training data for the classification task. After training the classifier with these seed data, unlabeled pixels are labeled by classifying these unlabeled data as test data. Our method consists of two major steps: (1) supervised classification of superpixels as initial segmentation and (2) pixel-wise classification as label refinement.

In supervised classification of superpixels, the goal of this step is to initially extract the target object from an image with the provided input seeds using supervised classification on the superpixelized image. The use of superpixels often gives advantages in computational efficiency by reducing the number of handled data and some positive effects of using clusters including robustness to inhomogeneity and usability of semantic information. In experiments, we use SLIC superpixels [1] as a method of superpixelization which provides accurate superpixels at the fast speed. Supervised classification mainly consists of two sub-steps: (1) feature vector generation, and (2) training and classification. In feature vector generation, image components according to each superpixel, including color and pixel location, are used to generate the superpixel features. For an image I with the arrays of color components, $\mathbf{RGB}$ in RGB color space, which are averaged for each superpixel, and the arrays of $x-$ and $y-$coordinates of centers of superpixels, $\mathbf{x}, \mathbf{y}$, the set of features $\mathbf{F}$ is defined as

$$\mathbf{F} = \{\mathbf{f} | \mathbf{f} \in \{\mathbf{R}, \mathbf{G}, \mathbf{B}, \mathbf{x}, \mathbf{y}\}\}. \tag{1}$$

With the generated feature vector, we train and classify the binary classifier with the seed superpixels as training data, and the rest as test data. In experiments, we use support vector machines (SVM) with radial basis function (RBF) kernel for supervised classification.



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

Figure 1. Experimental results; (a) Original images, (b) provided seeds, (c)-(e) object-labeled images segmented by (c) Random Walks (RW) [3], (d) the proposed method with color features-only (RGB), and (e) the proposed method with color and location features (RGBxy).

Though the superpixelization is an edge-preserving process, the object label resulted from supervised classification of superpixels often has outliers at its boundaries. To refine the label by correcting the outliers at the boundaries, we perform label refinement with pixel-wise classification. First, we generate the tri-map dividing the image into object, background, and unknown regions, from the segmented label by generating the narrow band around the boundaries between object and background labels. With the generated tri-map, we repeat the process of supervised classification, but in this time, on pixels. In experiments, we use k-nearest neighbors (kNN) for pixel-wise classification.

3. Results and Conclusion

We evaluated the performance of our method and compared it to those of other state-of-the-other interactive segmentation models. For validation, we use the public data called GrabCut [5] and Gulshan [4] data sets. GrabCut data includes 50 natural images with lasso-labeled seeds and ground truth labels, while Gulshan data includes 151 natural images with scribble-labeled seeds and ground truth labels. Fig. 1 shows the experimental data and their results. As shown in Fig. 1 (b), object and background seeds are represented as white and dark gray regions in GrabCut data at the top two rows, and as white and red lines in Gulshan data at the bottom three rows, respectively.

Table 1. Label error rates (%) of the proposed and comparative methods in two experimental data sets.

Data sets	GrabCut [5]	Gulshan [4]
RW [3]	3.3	10.8
RGB	9.2	20.9
RGBxy	**2.6**	**7.6**

Fig. 1 (c)-(e) show the experimental results segmented by (c) random walks (RW) [3], (d) the proposed method with color feature-only (RGB), and (e) the proposed method with color and location features (RGBxy). As shown in (c), RW inaccurately extracts the object especially when the amount of seeds are very limited such as in Gulshan data. In the proposed method, RGB includes the outlier region as shown in (d) while RGBxy shows the finest performance by excluding these outliers as shown in (e).

To validate and compare the performances of our method numerically, we computed the error rates of labeled pixels, which is a ratio between the number of mis-labeled pixels and the number of initially unknown pixels. Table 1 shows the label errors computed for the proposed and comparative methods in GrabCut and Gulshan data sets. As shown in the table, our method with color and location features scores the lowest error rate compared with the state-of-the-art RW method.

In this paper, the method of interactive image segmentation based on supervised classification with superpixels was proposed. In experiments, our method showed competitive performance compared to the state-of-the-art method even with the very simple choice of features and classifiers. Future works will focus on evolving the potential of our method by developing novel features suitable for our method and enhancing the method in numerous ways.

References

[1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(11):2274–2282, Nov 2012. 1

[2] Y. Boykov and M.-P. Jolly. Interactive graph cuts for optimal boundary amp; region segmentation of objects in n-d images. In *Proc. ICCV*, volume 1, pages 105–112 vol.1, 2001. 1

[3] L. Grady. Random walks for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(11):1768–1783, Nov 2006. 1, 2

[4] V. Gulshan, C. Rother, A. Criminisi, A. Blake, and A. Zisserman. Geodesic star convexity for interactive image segmentation. In *Proc. CVPR*, 2010. 2

[5] C. Rother, V. Kolmogorov, and A. Blake. Grabcut: Interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.*, 23(3):309–314, Aug. 2004. 2