# Object Segmentation by Edges Features of Graph Cuts

Weiwei Du, Yuki Masumoto, Nobuyuki Nakamori
Dept. Information Science
Hashigamicho, Matugasaki, Kyoto, 606-8585 Japan
e-mail: {duweiwei, m2622037, nakamori}@kit.ac.jp

*Abstract*—**This paper proposes a simple graph cuts algorithm based edges features to object segmentation problems. Graph cuts are used to find the global optimum of a cost function based on boundary and region of an image. Gaussian Mixture Models (GMMs) are built based on the seeds which are given by user to the object and background in an image. The contribution of this paper is to add edges features to GMMs. The proposal can segment an object region having noisy edges and colors similarity between the object and background. Experimental results illustrate the validity of the proposal.**

*Keywords-object segmentations; edges features; graph cuts; Gaussian Mixture Models*

## I. INTRODUCTION

Object Segmentation in static images is one of the most fundamental tasks in image content analysis, object recognition, image matting and so on. Many algorithms of object segmentation have been proposed. For example, an object can be segmented by obtaining a $\alpha$ which is expressed at $I= F\alpha +B (1- \alpha )$. $F, B$ and $\alpha$ are foreground colors, background colors, opacity respectively. However, $F, B$ and $\alpha$ are unknown, object segmentation is the highly ill-posed problem. In order to solve the above problem, some object segmentation algorithms have proposed such as the trimap algorithm [1][2], membership propagation algorithm [3]. However, they are time-consuming.

Alternatively, from minimize certain energy functions view, graph cuts algorithms are widely explored such as early organ segmentation [4], interactive graph cuts [5], and recently efficient N-D image segmentation [6], image segmentation using multi-scale smoothing [8] and graph cuts segmentation using local texture features [9]. The goal of these algorithms is to obtain a minimize certain energy function which is defined in terms of boundary and region in an image. Their basic idea of the algorithms is Min-Cut/Max-Flow in graph theory. The minimize energy can be obtained based on adding or removing any constraints by user.

Early graph cuts algorithms developed with an image are difficult to segment an image with complex noisy edges, because these noisy edges interfere with the values among neighboring pixels. Thus, [8] proposed a coarse-to-fine approach to detect the true boundaries using graph cuts. However, [8] cannot segment the kind of images which have some analogue colors between the object and background. Graph cuts segmentation using local texture features of multiresolution [9] can solve the above problem, to some extent, but it is not easy to get the texture features from an image.

This paper proposes a simple graph cuts algorithm based on edges features to object segmentation problems. Graph cuts are used to find the global optimum of a cost function based on boundary and region of an image. Gaussian Mixture Models (GMMs) are built based on the seeds which are given by user to the background and foreground in an image. The contribution of this paper is to add edges features to GMMs. The proposal can segment an object region having noisy edges and colors similarity between the background and foreground, and improve precision. Experimental results illustrate the validity of the proposal.

The paper is organized as following: Section two introduces theory of graph cuts in an image. Section three interprets the proposal of this paper. Section four expounds the procedures of the approach. Section five draws a conclusion and gives a future work.

## II. GRAPH CUTS FOR AN IMAGE SEGMENTATION

Boykov et al. [4][5][6] apply the theory of graph cuts algorithm to image segmentation. An undirected graph $G=(V,E)$ is defined as a set of nodes $V$ and a set of undirected edges $E$ that connect these nodes. A simple graph corresponding to an image is shown in Figure 1. In the graph, there are two terminals which are called as "a source" and "a sink". In the image, the source is considered as the object and a sink is considered as the background. The image can be segmented when every pixel corresponds to the source or the sink. However, all pixels of an image cannot be labeled, we must judge the unlabeled pixels which belong to the object or the background based on labeled pixels. We call the edges of the labels corresponding to a source or a sink *t-link*. The edges of neighboring pixels are called *n-link*. We obtain the segmentation boundary between the object and the background by computing *n-link* and *t-link*. Obtaining the segmentation boundary means finding the minimum cost cut on the graph. It is note that locations with high intensity gradient correspond to cheap *n-link*. Thus, they are attractive choices for optimal segmentation boundary. The minimum cut can be computed exactly in polynomial time using well known algorithms such as max-flow [10] or push-relabeled [11].
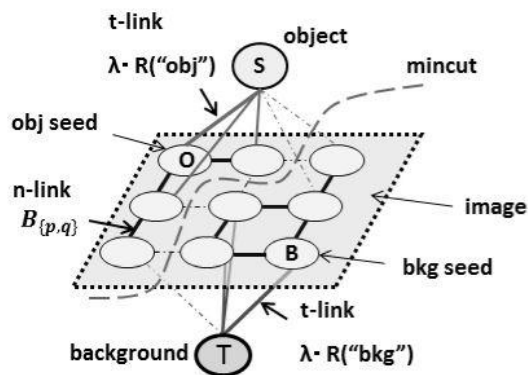
Figure 1. Graph representing a $3 \times 3$ image

Actually, image segmentation is considered as a binary labeling problem. The nodes are pixels $p$ on the image $P$ and the edges have adjacency relationships with four or eight connections between neighboring pixels $q \in N$. $N$ is a set of neighboring pixels. The labeling problem is assigned a unique label $A$ to each node. $A = (A_1, A_2, ..., A_p, ... A_{|p|})$ can be obtained by minimizing the energy $E(A)$ in Eq.(1). A is a binary vector i.e. $A_p \in \{"obj", "bkg"\}$. "$obj$" and $O$ are represented object while "$bkg$" and $B$ are represented background in an image. $P$ is the number of pixels on the image.

$$E(A) = \lambda \cdot R(A) + B(A) \qquad (1)$$

Where

$$R(A) = \sum_{p \in P} R_p(A_p) \qquad (2)$$

$$B(A) = \sum_{\{p,q\} \in N} B_{\{p,q\}} \cdot \delta(A_p, A_q) \qquad (3)$$

$$\delta(A_p, A_q) = \begin{cases} 1 & A_p \neq A_q \\ 0 & otherwise \end{cases}$$

The coefficient $\lambda \geq 0$ in Eq. (1) specifies the relative importance of the region properties term $R(A)$ shown at Eq. (2), to the boundary properties term $B(A)$ shown at Eq. (3). The term $R(.)$ reflects how the intensity of pixel $p$ fits into a known intensity model of object and background. The term $B(A)$ comprises the boundary properties of segmentation. $B(A)$ is interpreted as a penalty for discontinuity between pixels $p$ and $q$. $B_{\{p,q\}}$ is normally large when $p$ and $q$ are similar.

Table I

EDGE COST

| edge | | cost | for |
|---|---|---|---|
| n-link | $\{p, q\}$ | $B\{p,q\}$ | $\{p,q\} \in N$ |
| t-link | $\{p, S\}$ | $\lambda \cdot R_p("bkg")$ | $p \in P, p \notin O \cup B$ |
| | | $K$ | $p \in O$ |
| | | $0$ | $p \in B$ |
| | $\{p, T\}$ | $\lambda \cdot R_p("obj")$ | $p \in P, p \notin O \cup B$ |
| | | $0$ | $p \in O$ |
| | | $K$ | $p \in B$ |

Table I lists the edge costs of the graph. The region term and boundary term in Table I are calculated by:

$$\begin{cases} R_p("obj") = -\ln \Pr(O \mid C_p) \\ R_p("bkg") = -\ln \Pr(B \mid C_p) \end{cases} \qquad (5)$$

$$B_{\{p,q\}} \propto \exp(\frac{(I_p - I_q)^2}{2\sigma^2}) \cdot \frac{1}{dist(p,q)} \qquad (6)$$

$$K = 1 + \max_{p \in P} \sum_{q:\{p,q\} \in N} B_{\{p,q\}} \qquad (7)$$

$I_p$ is the brightness values and $C_p$ is *RGB* color features at pixel $p$. In Eq. (5), the likelihood is computed based on Gaussian Mixture Models. The boundary between the object and the background is found by searching for the minimum cost [7] on graph $G$.

III. OUR APPROACH

The paper proposes a simple graph cuts algorithm based edges features to object segmentation problems. Graph cuts are used to find the global optimum of a cost function based on boundary and region of an image. Smoothing the image using downsampling and upsampling is to obtain the minimize energy while we do not need add or reduce seeds by user. The *n-link* can be computed after downsampling and upsampling of the image. Gaussian Mixture Models (GMMs) are built based on the seeds which are given by user to the background and foreground in an image. The *t-link* is computed by GMMs. The contribution of this paper is to add edges features to GMMs. The proposal can

segment an object region having noisy edges and colors similarity between the object and background.

### A. Smoothing image by downsampling and upsampling(n-link)

We use a max flow algorithm [10] to determine the minimum cut corresponding to the optimal segmentation. The max flow algorithm gradually increases the flow sent from the source $S$ to the sink $T$ along edges in $G$ given their capacities (weights). Upon termination the maximum flow saturates the graph. The saturated edges correspond to the minimum cost cut on $G$ giving us an optimal segmentation. In original image segmentation by graph cuts [4], User added or reduced seeds for changing the capacities of graph. In this paper, we change the capacities of graph by the coarse-to-fine level which is shown at the Figure 2. It is the same as Nagahashi [8].
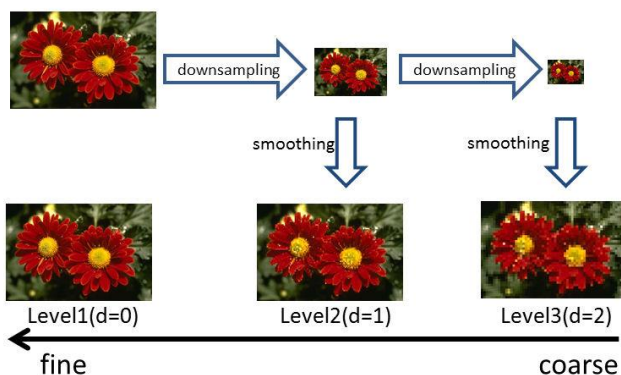


Figure 2. Smoothing the image by downsampling

### B. Edges features by the monochrome images

As we all know, the traditional graph cuts algorithm is difficult to handle the kind of images which have some noise or analogue colors between the object and background. That is because GMMs just used the colors information. Thus, we add the edges features to GMMs. As a sobel filter has the characteristic to control the noise, we adopt it for obtaining edges features of the image. The monochrome images are used with formula Y=0.299R+0.587G+0.114B in order to compute the sobel filter conveniently. Y stands for luminance. We need obtain edges features of every image which is shown in Figure3.
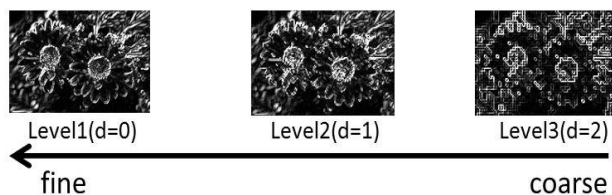


Figure 3. Edges features in the monochrome image

### C. Object segmentation by graph cuts

4 dimensional features $X_p = \{C_p, E_p\}$ are derived from R, G, B color features $C_p$ and edges features $E_p$. In Eq. (5), t-link edge costs are transformed to the posterior probability to achieve greater further accuracy as follows:

$$\begin{cases} R_p{'}("obj") = -\ln \Pr(O \mid X_p) \\ R_p{'}("bkg") = -\ln \Pr(B \mid X_p) \end{cases} \quad (8)$$

The posterior probability is proportional to the product of the prior probability (Pr(O), Pr(B)) and the features likelihood according to Bayes' theorem as follows:

$$\begin{cases} \Pr(O \mid X_p) = \dfrac{\Pr(X_p \mid O)\Pr(O)}{\Pr(X_p)} \\ \Pr(B \mid X_p) = \dfrac{\Pr(X_p \mid B)\Pr(B)}{\Pr(X_p)} \end{cases} \quad (9)$$

$$\begin{cases} \Pr(O) = \begin{cases} d_{obj} & d_{obj} \le d_{bkg} \\ 1 - d_{obj} & otherwise \end{cases} \\ \Pr(B) = 1 - \Pr(B) \end{cases} \quad (10)$$

The feature likelihoods $\Pr(X_p \mid O)$, $\Pr(X_p \mid B)$ are derived using Gaussian Mixture Models. The GMMs is obtained by:

$$\Pr(X_p \mid \cdot) = \sum_{i=1}^{K} \alpha_i p_i(X_p \mid \mu_i, \Sigma_i) \quad (11)$$

$$p(X_p \mid \mu, \Sigma) = \frac{1}{(2\pi)^{4/2}|\Sigma|^{1/2}} \cdot e^{(\frac{1}{2}(X_p - \mu)^T \Sigma^{-1}(X_p - \mu))}$$

$$(12)$$

We use the EM algorithm to fit the GMMs.

We initialize Pr(O)=Pr(B)=0.5 in d level, As we consider a pixel of the image just has two classes which are the object or background. The likelihood $\Pr(X_p \mid O)$, $\Pr(X_p \mid B)$ are derived using Gaussian Mixture Models with seeds of the object and background by user. t-links of the object and background are computed as the (d-1)th level image in Eq. (8). The (d-1)th level image can be segmented based in Eq. (13). The prior probabilities Pr(O) and Pr(B) of (d-2)th level are computed using the distance transform of

the segmentation result of (*d*-1)th level based on Eq. (10). The distance from the boundary is normalized from 0.5 to 1.0. $d_{obj}$ is defined as the normalized distance to the object, and $d_{bkg}$ is defined as the normalized distance to the background. $\Pr(X_p|O)$, $\Pr(X_p|B)$ of (*d*-2)th level are computed based on (*d*-1)th level image segmentation. It is shown in Figure 4. This processing is repeated until *d*=0.

$$
\begin{cases}
p \in O & \left| \lambda \cdot R_p^{'}("obj") - \sum_{q:\{p,q\} \in N} B_{\{p,q\}} \right| \\
& < \left| \lambda \cdot R_p^{'}("bkg") - \sum_{q:\{p,q\}} B_{\{p,q\}} \right| \\
p \in B & \left| \lambda \cdot R_p^{'}("obj") - \sum_{q:\{p,q\} \in N} B_{\{p,q\}} \right| \\
& > \left| \lambda \cdot R_p^{'}("bkg") - \sum_{q:\{p,q\}} B_{\{p,q\}} \right|
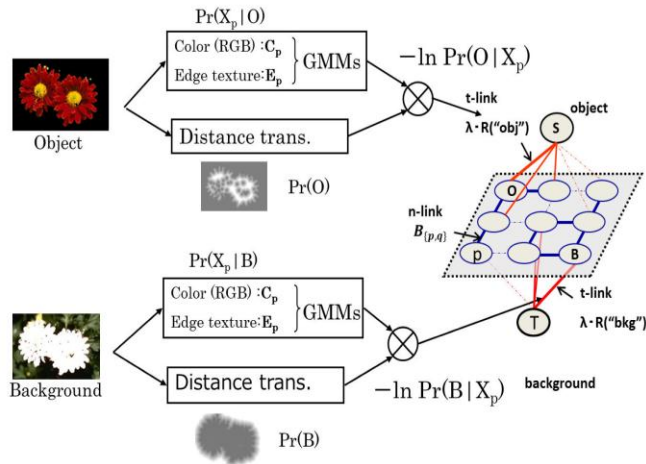\end{cases}
\quad (13)
$$



Figure 4. The posterior probability for {*p, S*} and {*p, T*}

## IV. STEPS OF OUR ALGORITHM

Figure 5 shows the process of the algorithm. It is carried out according to the following procedure.

  *1) Degrade a color image to the low resolution image with a downsampling method.*
  *2) Give some seeds the dth level color image.*

  *3) Obtain Y value of the low resolution monochrome image with the formula Y=0.299R+0.587G+0.114B.*
  *4) Extract the edges features of the monochrome image.*
  *5) Obtain t-link using the prior probability GMMs.*
  *6) Obtain n-link by the dth level color image.*
  *7) Obtain a segmentation image by graph cuts.*

Repeat from step 3 to step 6 until the original image is obtained. After that, go ahead to step 7 in order to get a segmentation image.
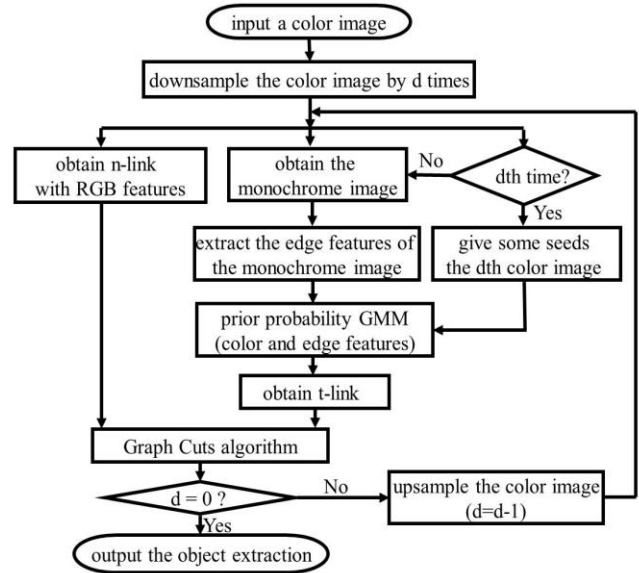


Figure 5. Flowchart on the process of our proposal

## V. EXPERIMENT

The images of segmentation experiments come from the Grab Cuts Database [12]. The image database has the original images and mask images of humans, animals, landscapes and so on. User gives seeds to the original images and the differences between mask images given in the database and output images are computed as error rate as shown in Figure 11. We compare interactive graph [5] (method1), image segmentation using multi-scale smoothing [8] (method2), graph cuts segmentation by texture features [9] (method3) and our proposal. The segmentation error rate is defines as:

$$
Error[\%] = \left( \frac{un\det ected \quad pixels \quad in \quad background}{image \quad size} + \frac{undertected \quad pixels \quad in \quad object}{image \quad size} \right) \times 100\%
$$

(14)

## A. An image with noise and analogue colors between the foreground and background

We give two images for describing our approach. Figure 6 is a difference colors image between the object and background by their histograms which are shown at Figure 8. Figure 7 is an analogue colors image between the object and background by their histograms which are shown at Figure 9. The horizontal axis of figure 8 and 9 is bins from 0 to 255 and the vertical one is the average values of R, G and B. Figure 6 and Figure 7 give the accurate objects and backgrounds based on [12]. We consider that figure 7 has an analogue colors image between the object and background, because they have analogue histograms between the object and background like a Gaussian distribution based on Figure 9. Conversely, we obtain a failure result referring to figure 12, because they have a difference colors image between the object and background based on Figure 8.

Finally, we add Gaussian noises to Figure 7. The experimental result is shown at Figure 10. The error rate of mehtod1 without edges features is 6.78% while the error rate of our approach with edges features is 2.06%. Our approach is effective to the kind of images which have noise and an analogue colors image between the foreground and background.



an original kangaroo     background of an kangaroo     foreground of an kangaroo

Figure 6. A different colors image between the foreground and background



an original book     background of a book     foreground of a book

Figure 7. An analogue colors image between the foreground and background



the histogram of an kangaroo's background     the histogram of an kangaroo's foreground
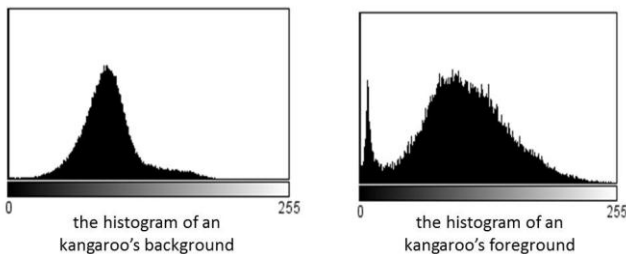
Figure 8. Histogram of the kangaroo image between the foreground and background



the histogram of a book's background     the histogram of a book's foreground
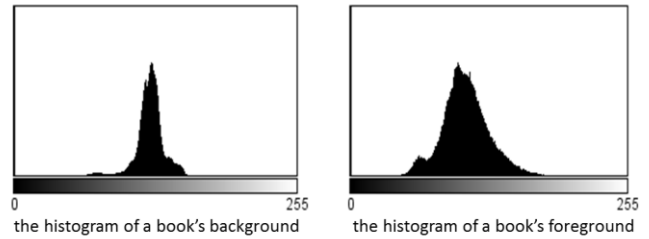
Figure 9. Histogram of the book between the foreground and background



an original book with noises     the result of method1 without edges features     the result of our approach with edges features
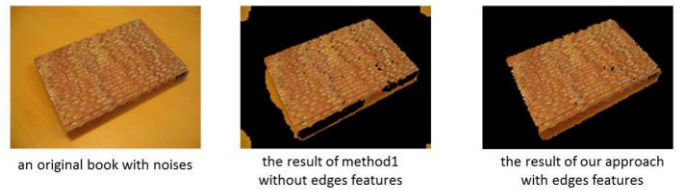
Figure 10. The results of the book image with noises

## B. Comparison with other methods

Our approach obtains the smallest error rate comparison with other methods. The results are shown in Figure 11. However, our proposal does not always obtain the smallest error rate to all images such as Figure 12. Generally speaking, our approach is always prior to method1 and method2. When the image has heavy texture change, our approach is inferior to method3.

## VI. CONCLUSION AND FUTURE WORK

We propose a simple graph cuts algorithm for object segmentation based on edge features of an image. We just add the edges features in GMMs in order to compute *t-link*. The proposal can improve the segmentation error rate compared to the conventional methods to the kind of the images which have an object region having noisy edges and colors similarity between the object and background. Experiments' results also certified the effectiveness of the approach. However, we must find the appropriate the values of $\sigma$ and $\lambda$ in order to apply our proposal. Therefore, future work is required adding texture feature for segmentation and the appropriate parameters $\sigma$ and $\lambda$ can be obtained automatically.

### REFERENCES

[1] Ruzon, M. and Tomasi, C.: "Alpha estimation in natural images," Proc. CVPR, pp. 18-25, 2000.

[2] Chunang, Y. Y., Curless, D., Salesin, D., and Szeliski, R. : "A bayesian approach to digital matting," Proc. CVPR, pp. 264-271, 2000.

[3] W. Du and K. Urahama : "Natural image matting with membership propagation," IPSJ Trans. CVA, vol. 1, no. 1, pp. 3-11, 2009.

[4] Y.Boykov and M.-P. Jollu.:"Interactive organ segmentation using graph cuts," In Medical Image Computing and Computer-Assisted Intervention, pp. 276-286, 2000.

[5] Y.Boykov, Y., Jolly, and M.-P.: "Interactive graph cuts for optimal boundary and region segmentation of object in N-D image segmentation," Internation Conference on Computer Vision, vol. I, pp. 105-112, 2001.

[6] Y.Boykov and G. Funka-Lea: "Graph cuts and efficient N-D image segmentation," Int. J. Comput. Vis. vol. 70, no. 2, pp. 109-131, 2006.

[7] Y.Boykov and V. Kolmorov: "An experimental comparison of mincut/max-flow algorithms for energy minimization in vision," IEEE Trands. Pattern Anal. Mach. Intell. vol. 26, no. 9, pp. 1124-1137, Sep. 2004.

[8] T. Nagahashi, H.Fujiyoshi, and T.Kanade: "Image segmentation using iterated graph cuts based on multi-scale smoothing," Asian Conference on Computer Vision, Part II, LNCS 4844, pp. 806-816, 2007.

[9] K. Fukuda, T. Takiguchi, and Y. Araki: "Graph cuts segmentation by using local texture features of multiresolution analysis," IEICE TRANs. INF. and SYST. vol. E92-D, no. 7, pp. 1453-1460, 2009.

[10] L. Ford and D. Fulkerson.: "Flows on networks," Princeton University Press, 1962.

[11] A. Goldberg and R. Tarjan. : "A new approach to the maximum flow problem," Journal of the Association for Computing Machinery, 35(4): 921-940, October 1988.

[12] http://research.microsoft.com/en-us/um/cambridge/projects/ visionimagevideoediting/segmentation/grabcut.htm/. Retrieved: September, 2012
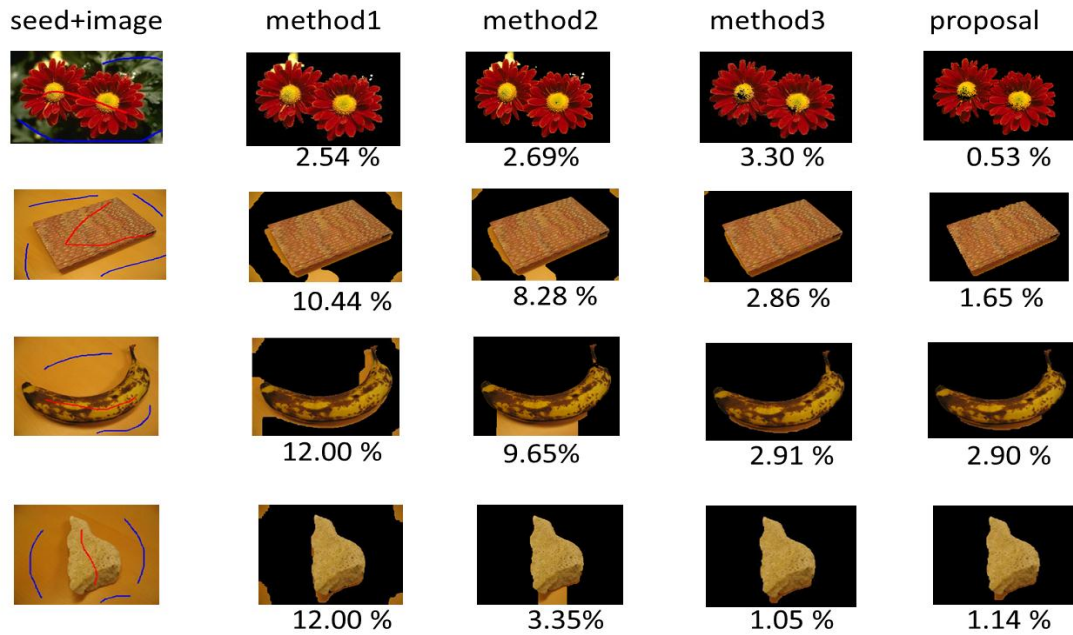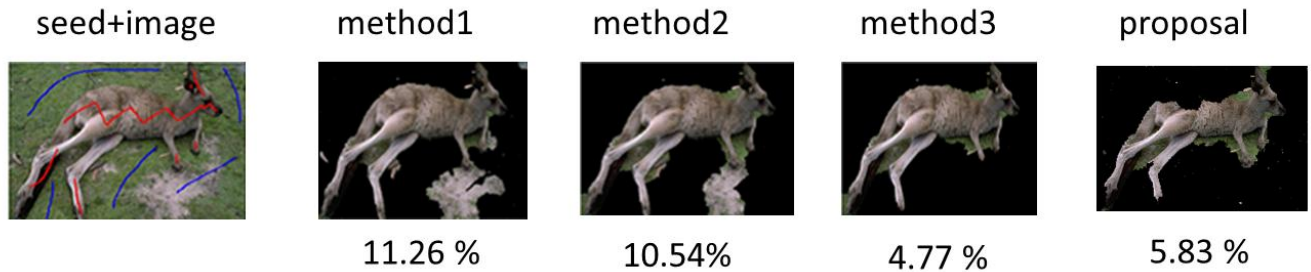
Figure 11. Examples of segmentation results



Figure 12. A failure example of segmentation result