Improved Minimum Spanning Tree based Image Segmentation with Guided Matting

Weixing Wang[1*], Angyan Tu[2*], Fredrik Bergholm[3]
[1]School of Information Engineering, Chang'an University, Xi'an, China, 710064
[e-mail: wxwang@chd.edu.cn]
[2]Department of Computer Science and Eng., Shaoxing University, Shaoxing, China, 312000
[e-mail: lucytu@shu.edu.cn]
[3]Royal Institute of Technology, Stockholm, Sweden, 10044
[e-mail: fredrikb@kth.se]
*Corresponding author: Weixing Wang and AngyanTu

Abstract

In image segmentation, for the condition that objects (targets) and background in an image are intertwined or their common boundaries are vague as well as their textures are similar, and the targets in images are greatly variable, the deep learning might be difficult to use. Hence, a new method based on graph theory and guided feathering is proposed. First, it uses a guided feathering algorithm to initially separate the objects from background roughly, then, the image is separated into two different images: foreground image and background image, subsequently, the two images are segmented accurately by using the improved graph-based algorithm respectively, and finally, the two segmented images are merged together as the final segmentation result. For the graph-based new algorithm, it is improved based on MST in three main aspects: (1) the differences between the functions of intra-regional and inter-regional; (2) the function of edge weight; and (3) re-merge mechanism after segmentation in graph mapping. Compared to the traditional algorithms such as region merging, ordinary MST and thresholding, the studied algorithm has the better segmentation accuracy and effect, therefore it has the significant superiority.

1. Introduction

In graph-based image segmentation methods, the boundary of a connected region is equal to the edge of the region, therefore, it is a problem that the extracted edge is closed but without additional conditions or criteria to judge. Since1971, Zahn first applied graph theory to image segmentation and data clustering [1], the research for the graph-based algorithms became a hot issue in the world for its good mathematical foundation and application extensions. After then, a variety of graph-based methods were developed. The methods mainly include graph cut based [2-4], interactive based [5], minimum spanning tree (MST) based [6-7], shortest paths based [8] and pyramid-based algorithms [9].

The graph cut based algorithms have been studied more in the last 20 years. The earliest algorithm was introduced by Wu and Leahy in 1993 [2], and it cuts edges that have the weak similarity between pixels. Their algorithm is prone to produce small regions, thence Shi and Malik*,* et al. proposed a normalized cut algorithm [10] that considers the global information of an image to maximize the differences between the regions as well as the similarity within the regions on the basis of Wu and Leahy's, and made a slight improvement on the previous algorithm [11]. Boykov, et al. explored an optimization framework of the energy function based on graph cut and applied it for ND image segmentation to reach the global optimum. It is efficient in practical applications, robustness in computing, and is similar to the function of the Mumford-Shah with the framework based on the area's properties.

In 2006, Sharon, et al. published their research result in "Nature", described a top-down hierarchical segmentation algorithm [12], calculated the coupling edge weight in the establishment of the subdivision graph, selected several seeds to coarse segment literately, obtained salience map, and extracted significant targets in an image. The segmentation result is accurate and efficient. Ma*,* et al. used the graph spectrum theory and Watershed idea to segment Synthetic Aperture Radar (SAR) images [13], and improved the efficiency in handling large resolution images. Chen, et al. presented a *k* clustering and graph cuts-based image segmentation algorithm [14], achieved automatic segmentation of cardiac dual-source CT images and picked up the structure of heart exactly. Hou studied the automatic segmentation of a static image and moving object image by combining graph cuts with level set [15]. Egger, et al. developed a scale fixed template-based paradigm [16] on "Nature" in 2012, for the complex images where the gray levels in a target area are similar to the background's, brought in the concept of the "template shape" to sample nodes in the graph non-uniformly and non-equidistant, furthermore, applied the suggested algorithm to the brain tumor image processing in 2D and 3D respectively.

In addition to the representative algorithms above, there are live wire algorithms. Wieclawak and Pietka modified live wire algorithm [17], integrated Wavelet transform with Fuzzy *C* means clustering to build a cost function, and it uses the graph theory into the medical images of the computed tomography (CT) and magnetic resonance (MR). Except for that, the random walk and isoperimetric algorithm is rose in recent years. Gopalakrishnan, et al. abstracted significant targets through a random walk [18] in graph, got the global attributes based on the detected seed nodes by the Markov random walk, computed the local properties by the sparse *k*-regular graph, determined seed nodes of background and foreground via semi-supervised learning step by step, achieved good accomplishment, too. Yi, et al. combined the random walk algorithm with Mean Shift [19], weakened the shortcoming that the target contours are vulnerable to natural textures in background and improved the precision and efficiency of image segmentation. Wang, et al. modified the Isoperimetric algorithm to enhance noise immunity in images and increased work efficiency [20]. Unlike the above

algorithms, Felzenszwalb and Huttenlocher showed a "small region merging" segmentation criterion [21], that is, when internal differences of the area are greater than the differences between areas, then the two areas identified are homogeneous regions, and they are merged. Their graph-based algorithm is able to adapt to the image's characteristics to segment images to some extents and it works fast in practice, therefore, this kind of studies are in the light of this algorithm, and all the improvements for the algorithm are carried out based on the characteristics of a kind of special images.

The shortcomings of graph-based algorithms are that the bigger the parameters are set, the more details of the image are lost, while small parameters would prone to over-segmentation, the segmentation scale is difficult to correctly grasp, thereby, the foreground and background in an image partition wrong. To solve these problems, the segmentation procedure can be divided into two steps, the pre-segmentation or coarse segmentation can be considered firstly, and then the matting is one of the possible algorithms.

The matting algorithm can split the overall objects from a complex background in an image although it cannot detect the detailed information of the foreground, and make up the lack of graph-based algorithms in this regard [22-23]. The Bayesian matting algorithm was proposed by Chuang and Curless, they improved the Bayesian matting algorithm as most of pixels to be detected contain mixed colors which are the combination of foreground and background [24]. However, it is more appropriate in the image that the color distributions of the foreground and the background do not overlap. Levin suggested a closed-form solution to the natural image matting [25], it predicts the properties of the solution by analyzing the eigenvectors of a sparse matrix, and find the globally optimal alpha matte by solving a sparse linear system of equations. A guided feathering/matting algorithm [26] was made by He, and it is a fast and non-approximate linear time algorithm. However, the matting algorithms cannot be applied for image segmentation further for the prospects; hence it is restricted in some areas that need to extract the information of objects more concrete and more details [27].

Unlike artificial objects, in some situations, the natural objects are harder to detect by an ordinary image segmentation algorithm because the natural objects have rough surfaces, 3D geometry, different colors, variable sizes and irregular shapes, they are mixed with similar colors and shaped objective background (e.g., sky, water and clouds etc.) and the other non-individual objects such as woods or grass etc. If the background is simple and clear, such as the boat, sika deer and sparrow in **Fig. 1**, the targets are easily extracted by both the similarity-based and the discontinuity-based algorithms which can give the basic judgment information if the image can be well segmented by some intelligent algorithms such as Watershed, Dynamic thresholding [28], Level set [29], Minimum Spanning Tree [6-7], 3D image processing [30] and Neural network or Deep learning [31] based algorithms/methods, etc. The following three images are typical examples, three basic and typical algorithms are utilized for the image segmentation testing, and the image segmentation results are satisfactory. For the same targets in the complicated background, the targets are difficult to be extracted by both the similarity-based and the discontinuity-based algorithms and the more intelligent Clustering algorithm, as shown in **Fig. 2**.

In **Fig. 2**, the first image presents a lake with a boat, trees, woods, grass, clouds, blue sky, due to the effects of the objects, and the lake surface color varies from place to place. The similarity-based algorithm such as the Otsu algorithm can mostly separate the lake from the other objects: the grass and its reflection parts in water, the two trees are extracted, the woods are detected as red and dark colors, but the clouds mixed with the blue sky, the boat cannot be extracted clearly, just its reflection part in water is extracted almost. In contrast, the discontinuity-based algorithm such as Canny edge detector can detect boundaries of the boat

and clouds, but it cannot distinguish where is grass and where is woods. The more complicated image segmentation such as Clustering algorithm seems to be better, but it has obvious over-segmentation problem, e.g., it cuts the two trees into different small parts, and it also cuts the grass into different parts.

In the second image in **Fig. 2**, there is a sika deer in the grass, it seems to be a simple image, but it is difficult to extract the dear from the grass. For examples, the Otsu algorithm detects the grass into many different sized and shaped green and black regions, and the deer body also has spot and stripe black regions, that means that the dear cannot be completely taken out from the grass; the Canny edge detector only detects the deer head, butits body information is mixed with the grass; and the Clustering algorithm segments the dear into different small regions, which is over-segmentation problem, since some deer regions are mixed with grass regions, it is hard to do merging procedure.
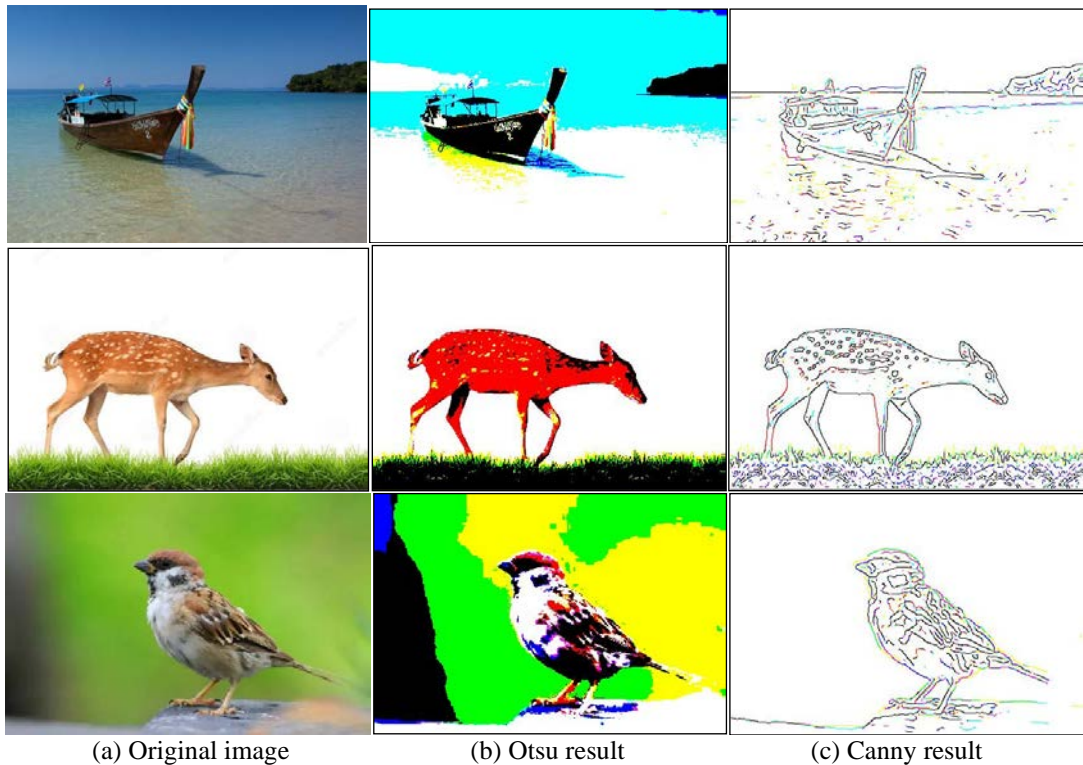
(a) Original image   (b) Otsu result   (c) Canny result
Fig. 1. Image segmentation by similarity-based and discontinuity-based algorithms.

The third image of **Fig. 2** is for a sparrow at a tree branch which is the second thick in the five tree branches, and some tree leaves are blurring. The Otsu algorithm detects most parts of 4 branches with black colors except for the thickest branch, and the bird body is clearly extracted except for the sparrow head that is connected to the background; the Canny edge detector just extract the bird and the second thick tree branch; and the Clustering algorithm also detects the second thick tree branch and the bird, even the branch includes different small regions, the region merging should be easy for image segmentation, but the other tree branches, especial the thickest branch cannot be extracted.

So, in order to make the good results of image segmentation for these complex natural images, a new algorithm is studied as the follows. It combines the matting algorithm and the improved graph-based algorithm, the former is for pre-segmentation (coarse segmentation) by

guided feathering, and the result is finalized for further accurate segmentation (fine segmentation) based on the improved graph-based algorithm.
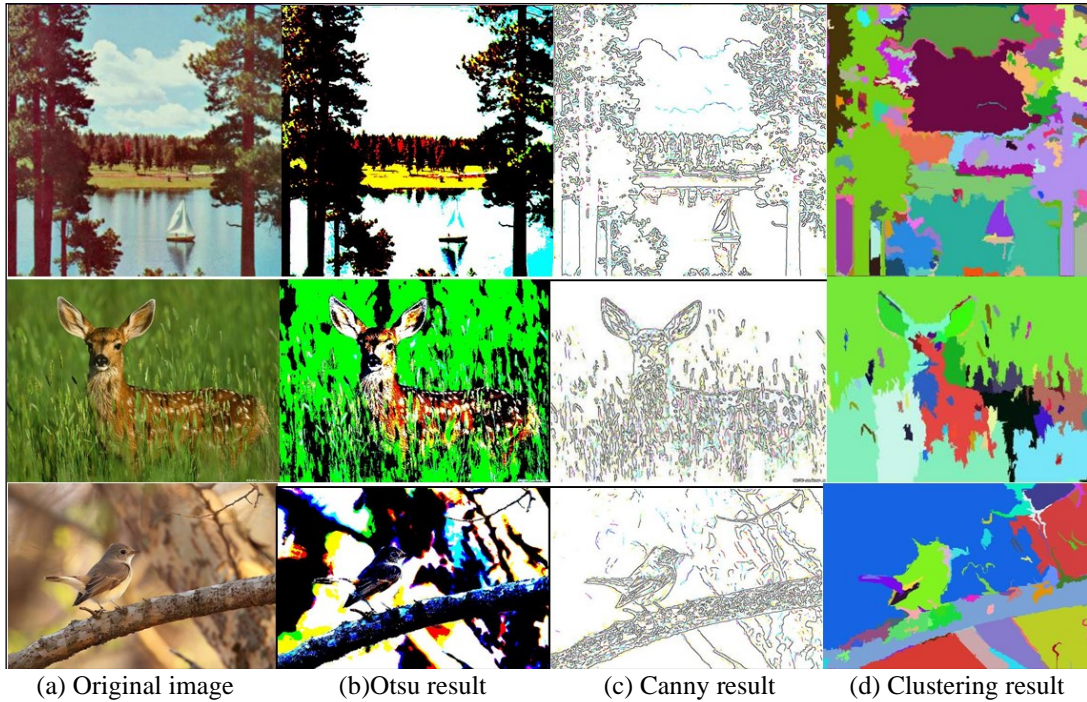

(a) Original image   (b)Otsu result   (c) Canny result   (d) Clustering result
Fig. 2. Traditional image segmentation algorithms for three complicated natural scene images.

2. Coarse Image Segmentation based on Guided Feathering

The matting algorithm is to extract objects from background overall in an image. For an input image I, a color combination equation is displayed as follows:

$$I = \alpha F + (1-\alpha) B \tag{1}$$

Where, F and B denote the foreground and background colors respectively, and α represents theopacity, shows the proportion of foreground colors.

According to the theory of guided matting [23, 26], α is presented as a linear function of image I in a small window ω_k:

$$\alpha_i \approx a_k I_i + b_k, \qquad \forall_i \in \omega_k \tag{2}$$

Where, $a_k = 1/(F_i - B_i)$, $b_k = -B_i /(F_i - B_i)$ are linear coefficients in ω_k. In order to seek a solution that minimizes the difference between output α and input p in a model as in Eq.(1), the linear coefficients (a_k, b_k) are determined by minimizing the following cost function:

$$E(a_k, b_k) = \sum_{i \in \omega_k} \left((a_k I_i + b_k - p_i)^2 + \varepsilon a_k^2 \right) \tag{3}$$

Where, ε is a regularization factor penalizing large ak. The cost function above is the linear ridge regression model, and its solution is obtained by:

$$a_k = \frac{\frac{1}{|\omega_k|} \sum_{i \in \omega_i} I_i p_i - \mu_k \overline{p}_k}{\sigma_k^2 + \varepsilon} \tag{4}$$

$$b_k = \overline{p}_k - a_k \mu_k \tag{5}$$

Where, μ_k, σ^2 and k are the mean and variance of intensities in ω_k, $|\omega_k|$ and $\overline{p}_k = 1/|\omega_k|\sum_i$ $_{\in \omega_k} p_i$ are the number of pixels and the mean of p in this window separately. Then the output α can be obtained by a_k and b_k.

However, a pixel i is involved in all the overlapping windows ω_k that covers i, so the value of α_i in Eq.(2) is not identical when it is computed in different windows. A simple strategy is to average all the possible values of α_i. So, the filtering output is computed by

$$\alpha_i = \frac{1}{|\omega_k|} \sum_{k|i \in \omega_k} \left(a_k I_i + b_k \right) = \overline{a}_i I_i + \overline{b}_i \tag{6}$$

Where, $\overline{a}_i = 1/|\omega_k|\sum_{k \in \omega_i} a_k$ and $\overline{b}_i = 1/|\omega_k|\sum_{k \in} \omega_k b_k$ are the average coefficients of all windows overlapping i. Putting Eq.(5) into the equation above, it is:

$$\alpha_i = \frac{1}{|\omega_k|} \sum_{k \in \omega_i} \left(a_k \left(I_i - \mu_k \right) + \overline{p}_k \right) \tag{7}$$

Then, due to the linear dependence between p and α, the matting kernel is obtained by $\partial \alpha_i / \partial p_j$:

$$\frac{\partial \alpha_i}{\partial p_j} = \frac{1}{|\omega_k|} \sum_{k \in \omega_i} \left(\frac{\partial a_k}{\partial p_j} \left(I_i - \mu_k \right) + \frac{\partial \overline{p}_k}{\partial p_j} \right) \tag{8}$$

and the new equation is:

$$\frac{\partial \overline{p}_k}{\partial p_j} = \frac{1}{|\omega_k|} \delta_{j \in \omega_k} = \frac{1}{|\omega_k|} \delta_{k \in \omega_j} \tag{9}$$

Where, $\delta_{k \in \omega_j\, j} = 1$ when j is in window ω_k, otherwise $\delta_{k \in \omega_j} = 0$. Additionally, referring to Eq.(4), $\partial \alpha_k / \partial p_j$ can be calculated as:

$$\frac{\partial a_k}{\partial p_j} = \frac{1}{\sigma_k^2 + \varepsilon} \left(\frac{1}{|\omega_k|} \sum_{i \in \omega_k} \frac{\partial p_i}{\partial p_j} I_i - \frac{\partial \overline{p}_k}{\partial p_j} \mu_k \right) == \frac{1}{\sigma_k^2 + \varepsilon} \left(\frac{1}{|\omega_k|} I_j - \frac{1}{|\omega_k|} \mu_k \right) \delta_{k \in \omega_j} \tag{10}$$

Therefore, $\partial \alpha_k / \partial p_j$ can be computed from Eq.(9) and Eq.(10):

$$\frac{\partial \alpha_i}{\partial p_j} = \frac{1}{|\omega_k|^2} \sum_{k \in \omega_i, k \in \omega_j} \left(1 + \frac{\left(I_i - \mu_k \right) \left(I_j - \mu_k \right)}{\sigma_k^2 + \varepsilon} \right) \tag{11}$$

As seen by the discussion above, α value range is:

$$0 \le \alpha \le 1 \tag{12}$$

Where, α needs to be rounded as required, that is, its value is either 0 or 1. With regarding to a graph-based algorithm, it is set not processing the pixels in that the gray value is 0 in this study, i.e., black points. Thus, for a region in that α value is between 0 and 1, both prospect matting result and background matting result will process this area based on the graph theory, the final sum will emerge the repeated superposition of foreground and background.

The image processing is in accordance with the above theory, the three distinctive natural

images (non-artificial target) are selected as shown in **Fig. 2**, the composition of the objects and background with textures are all complex. **Fig. 3** shows the comparison of the guided feathering and closed-form methods, **Fig. 3(b)** is the closed-form result. It is obvious that the proposed algorithm can retain the original color information of the image and can split objects out from background overall, including the parts staggered, such as the deer's body and bushes, or sailboat's reflection, etc.; whereas, the closed-form algorithm generates the color distortions in the matting process, such as the deer's body and the trunk turn red, the sky between the gaps of branches and leaves are replaced by brown. Furthermore, the main advantage of the guided feathering algorithm over other matting algorithms is that it naturally has an O(N) time non-approximate function. Thus, the guided matting algorithm has the significant advantages on the precision as shown **Fig. 3(b)**.

(a) Guided matting   (b) Closed-form
Fig. 3. Comparison between guided matting and closed-form on images in **Fig. 2**.

3. Fine Segmentation with Improved Graph-based Algorithm

On the basis of the above guided matting results, the further image segmentation is based on the graph theory. The graph-based algorithm proposed by Felzenszwalb is to map an image into a weighted graph G(V, E), then to segment the image by using Krusal minimum spanning tree algorithm based on the merger strategy, this algorithm involves three parameters: σ, k and min_size. Where, σ is a parameter of the Gaussian filter; k is a key parameter of the threshold

function, used to control the segmentation scale; and the *min_size* is a re-merge parameter, the two regions can be combined if in one of them, the size of the two adjacent domains is less than *min_size*. Although the effect of the graph-based algorithm is good, the structure is simple and the computation is efficient, but it still has a few shortcomings, so the following improvements for the two last parameters are given respectively.

3.1 Concrete improvements of graph-based method

(A) Ameliorate of difference of Intra-regional and Inter-regional functions in Minimum Spanning Tree (MST).

Due to Felzenszwalb only represented the degree of difference in the region with maximum edge weight, and thus vulnerable to the impact of the noise and isolated points, the segmentation results will deteriorate, too. The intra-regional difference is redefined as:

$$Int(C) = 1/N * \sum_{e \in MST(C,E)} w(e) \tag{13}$$

That is, the intra-regional difference is equal to the average of the edge weights of MST in the region, where N is the number of edges of MST, and the scilicet $N = |C|-1$. Although this may reduce the sensitivity to some extent, the two regions that can merge in the original image might not be combined with this function, but the segmentation scale can be controlled by adjusting the parameter k. Similarly, the inter-regional difference can be defined as:

$$Dif(C_1, C_2) = \frac{1}{2} \left(\min_{v_i \in C_1, v_j \in C_2, (v_i, v_j) \in E} w(v_i, v_j) + \max_{v_i \in C_1, v_j \in C_2, (v_i, v_j) \in E} w(v_i, v_j) \right) \tag{14}$$

(B) Ameliorate of edge weight function

The above improved algorithm is based on the segmentation criterion. The segmentation effect of the graph-based algorithm mainly depends on two aspects: the segmentation criterion and the weighting function. The weighting function of Felzenszwalb's algorithm [21] is only for the absolute difference of the gray scale values, without taking the spatial position of each pixel into consideration, generally, the greater the space distance of two pixels, the weaker the correlation, and the bigger the edge weight penalty. Therefore, the weighting function is rewritten as:

$$w(v_i, v_j) = \mu(v_i, v_j) |I(v_i) - I(v_j)| + d(v_i, v_j) \tag{15}$$

Where, $I(v_i)$ and $I(v_j)$ are the gray scale values of pixels v_i and v_j separately; $d(v_i, v_j)$ is defined as the Euclidean distance between v_i and v_j:

$$d(v_i, v_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{16}$$

In Eq.(16), (x_i, y_i) and (x_j, y_j) denote the coordinates of v_i and v_j respectively. $\mu(v_i, v_j)$ is called adjustment factor that regulates the gray scale value difference and the weight coefficient of the distance between two pixels, it is an adaptive two-dimensional Gaussian factor:

$$\mu(v_i, v_j) = \frac{1}{\sigma_i \sigma_j \sqrt{2\pi(1-r^2)}} e^{-\frac{1}{2(1-r^2)} \left[\frac{(i-\mu_i)^2}{\sigma_i^2} - \frac{2r(i-\mu_i)(j-\mu_j)}{\sigma_i \sigma_j} + \frac{(j-\mu_j)^2}{\sigma_j^2} \right]} \tag{17}$$

Where, μ_i and μ_j are expectations of gray scale values of the direction pixels; σ_i and σ_j are the standard deviations of gray scale values of the direction pixels. Furthermore:

$$r = \mathrm{cov}(i, j)\Big/\sqrt{x^2 + y^2} \qquad\qquad (18)$$

$$\mathrm{cov}(i, j) = E(ij) - E(i)E(j) \qquad\qquad (19)$$

(C) Ameliorate of re-merge mechanism after mapping graph segmentation

Except for the above two points, the image re-merging also affects the segmentation result, the desired result cannot be gained by simply adjusting k. The reason is that the greater of k, the larger of threshold value τ under Felzenszwalb's segmentation criterion, so the decision D is more likely to be true and it will be over-merging. As shown in **Fig. 4**, the bacterium and the background have different transitional zones, the shallow portion of the gray scale values is difficult to define boundary. It separates the bacterium from background. But in **Fig. 4(b)** while $k = 150$, the bacterium is over-segmented; while in **Fig. 4(c)**, $k = 250$, although the background segmented is preferable, but some pixels that clearly belong to bacterium are classified into the background.
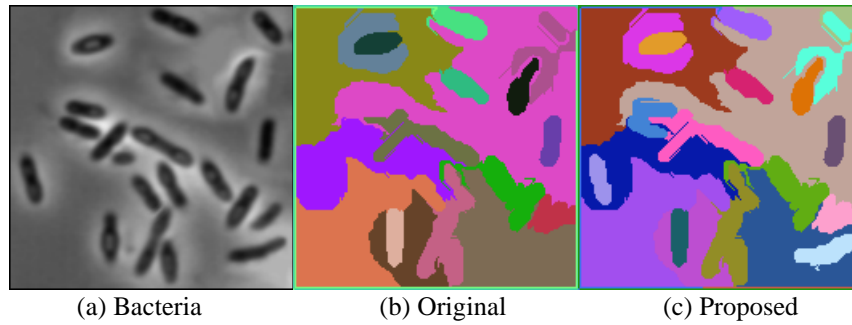

(a) Bacteria   (b) Original   (c) Proposed

Fig. 4. Comparison of merging mechanisms.

In order to overcome the above weakness, the re-merging mechanism can also be improved. Assume that q is the minimum size in a region; it can be set based on the target number in an image under normal circumstance. The merging mechanism is specifically described as follows: (1) Started from the division of S. For each edge $(v_i, v_j) \in E$, $v_i \in C_1$, $v_j \in C_2$, if $C_1 \neq C_2$ and $|C_1| < q$ and $|C_2| < q$, then C_1 is combined with C_2. The new segmentation S_{new} is got after all edges are processed; (2) For each edge of S_{new}, $(v_i, v_j) \in E$, $v_i \in C_1$, $v_j \in C_2$, if $C_1 \neq C_2$ and $|C_1| < q$ or $|C_2| < q$, C_1 is combined with C_2. The final image segmentation result is obtained by these two steps.

The above example illustrates why this merging mechanism can obtain relatively better results: the size of the pending segmented image is 3×3, and each of vertex of S is a region, $p = 4$, and it is from the up left to the low right for the processing edges.

As seen from **Fig. 5(a)**, step 1 is easier to obtain a large area, but some small regions still exist. After step 2, it can be ensured that the size of each region is greater than p. And at step 2, the probability of merging two different regions becomes lower; but it can be avoided to over-merging. Yet the original method jumps to step 2 directly, the image is only split into one area. Using the above method of improved re-merging mechanism to process the image in **Fig. 4(a)**, the result is better, when $k=250$, it is found that the phenomena of over-merging has been reduced in a certain of level.

(a) Proposed re-merge mechanism
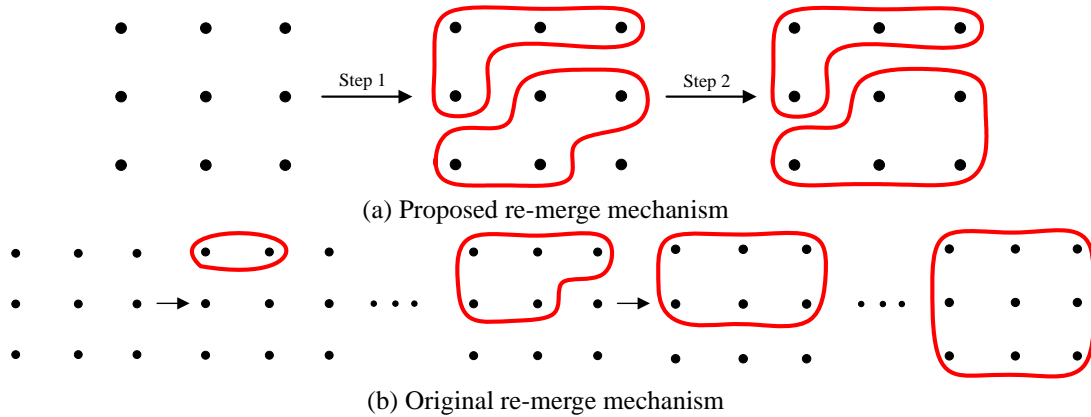

(b) Original re-merge mechanism

Fig. 5. Comparison of original re-merging mechanism and proposed re-merging mechanism.

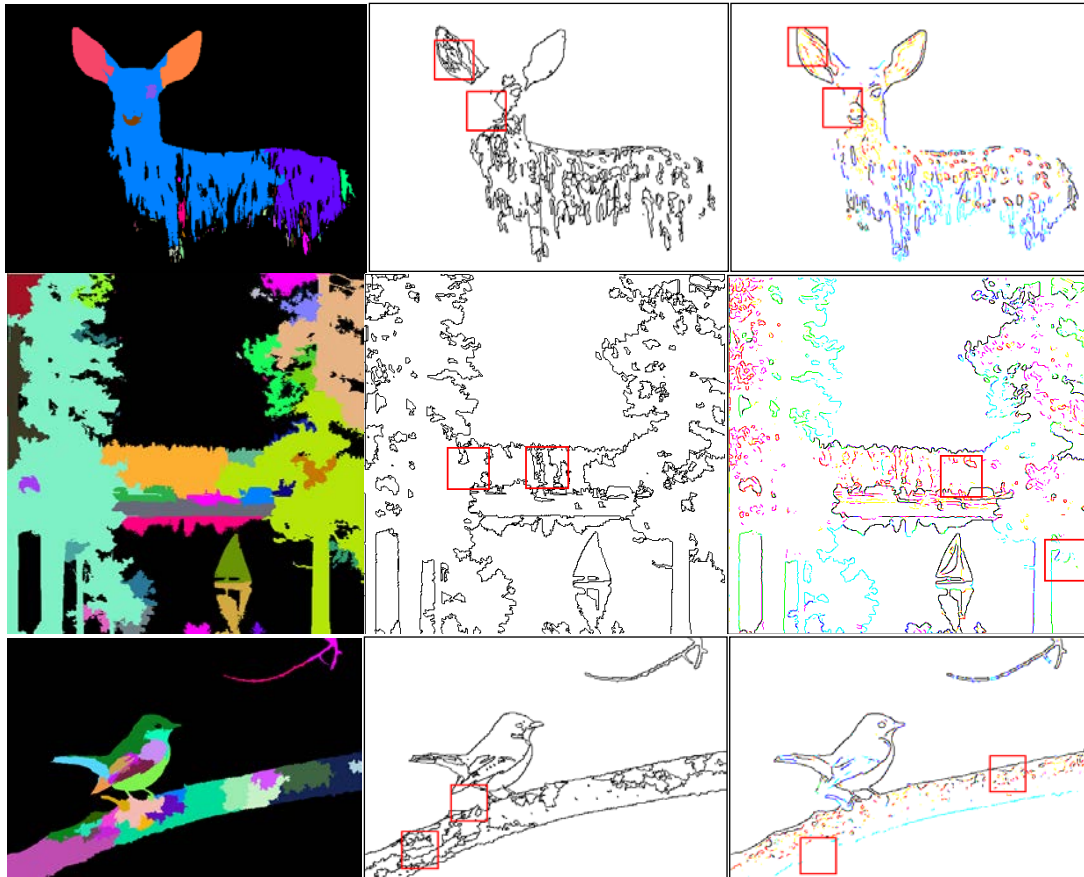Hence, the improved graph-based algorithm partitioning step is:

(1) Pre-processing: it uses a Gaussian filter to smooth the input image to remove noise;

(2) Mapping image to graph $G(V, E)$: it constructs a 8-*linked* weighted graph, $|V| = n$, $|E| = m$, the connection weights are set to $w(v_i, v_j)$;

(3) Sorting: all weights of E are arranged in a non-decreasing order to get a collection $\pi = (O_1, O_2, \dots, O_m)$;

(4) Initial State: assume that the initial segmentation is S^0, $S^0 = (v_1, v_2, \dots, v_n)$, i.e., each element (vertex) of V is a single region, repeat step 5 for $q = 1, 2, ..., m$;

(5) Looping: suppose S^{q-1} is the cut set after q-1 times merging, it establishes S^q from S^{q-1} as follows. Let $Cq\text{-}1\,i$ and $Cq\text{-}1\,j$ be the components of containing V_i and V_j respectively in S^{q-1}, if $Cq\text{-}1\,i \neq Cq\text{-}1\,j$, and $Dif(Cq\text{-}1\,i, Cq\text{-}1\,j) \leq \min(Int(Cq\text{-}1\,i) + \tau(Cq\text{-}1\,i), Int(Cq\text{-}1\,j) + \tau(Cq\text{-}1\,j))$, then merge the clusters $Cq\text{-}1\,i$ and $Cq\text{-}1\,j$ to get S^q, else, does not merge , $S^q = S^{q-1}$;

(6) Merging: return $S = S^m$, and then deal with S ground on the proposed re-merging mechanism;

(7) Labeling: assign the pixels to the same region with the same color in this step.

(8) Output: output image segmentation results.

3.2 Fine-segmentation on graph-based algorithm and comparing with others

For the guided matting results in **Fig. 2**, the images are enhanced firstly [32-33], then they are segmented by using the improved graph-based, the region merging [27, 34] based mean-shift model [35] and edge detection [36-37] algorithms respectively. The experimental results are shown in **Fig. 6**. As the following figures shown, without using pre-segmentation (closed form solution to matting) and the modified graph theory, the original images with complex edges are difficult to segment. Since the image noise is too much, the edge detection-based algorithms, for example, the Canny edge detector, cannot sketch out the contours and textures of the objects accurately.

With the comparison, the region merging algorithm has over-segmentation and/or under-segmentation problems in some places, all of the above images have the same problems more or less: the left half of the deer head is split out in the first row, and there are many disturbances in the body and the ear, i.e., over-segmentation; the trees and sailing in the second row are only depicted with the rough boundary, it cannot separate trees precisely; in the third row, the branches and tail of the bird are over-segmented too, and there are a lot of information

lost, such as the forked branches. In addition, this processing is slow; each color image of resolution 512x512 pixels needs more than 2 minutes for processing. The Canny segmentation result is not accurate enough and has the following problems: the edge is not continuous; the detection results contain more outliers (noised points). As the ear and body of the deer, the leaves of the second row and branches of the third row have the situation described above and the near trees are also not separated from the far distant trees.

(a) Improved graph-based   (b) Region merging   (c) Canny
Fig. 6. Segmentation results comparison on the basis of guided matting for the images in **Fig. 2**.

Table 1. Comparison of processing speeds of three algorithms (refer to **Fig. 6**)

	deer	sailboat	bird
Resolution	480×360	512×512	588×393
Proposed	6" 5	14" 2	7" 5
Region merging	2' 55" 3	2' 25" 4	1' 33" 5
Canny	1" 1	1" 3	1" 3

In the comparison, the effect of the improved graph-based algorithm is better, the part of deer body that mixed into the background grass is completely separated, the near trees are split out from the far distant trees, the bird's wings and abdomen in the third row also has the better segmentation result. Besides, it can be noticed that the graph-based algorithm can remove

outliers and noise effectively, and the high gradient magnitude value region can be combined into the same area.

The processing speeds of the above three algorithms are shown in **Table 1**, the mean shift preprocessing time is not taken into the consideration in the region merging algorithm, it is clear that the proposed algorithm has higher efficiency.

4. Experiments

In experiments, all the algorithms or operation functions used in this study are coded under the VC++ (version 6 or above) environment of Windows (version 8 and above) in a PC computer, and the hardware name and parameters are: DELL, CPU Intel Core i7-10700, Ram 8GB.
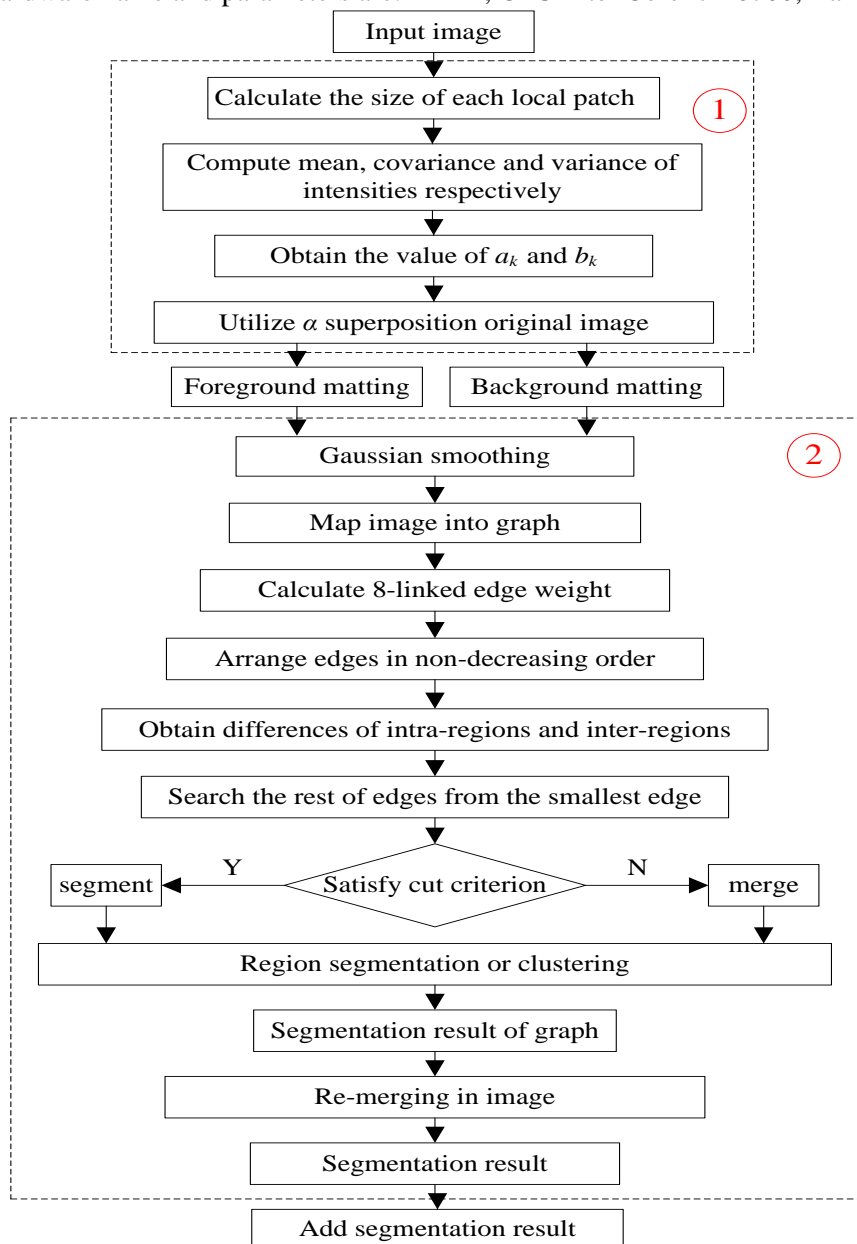

Fig. 7. Flow chart of proposed algorithm.

After the above analysis, the flow chart (including pre-segmentation and fine-segmentation) of the studied algorithm is shown in **Fig. 7**.

The red label 1 is for the guided matting part, and the number 2 is marked as the part of the improved graph-based algorithm. **Fig. 8** shows each step of the process according to **Fig. 7**. **Fig. 8(a)** is the original image, **Fig. 8(b)~(d)** are the tri-map and matting results of the guided matting part respectively, **Fig. 8(e)~(f)** are the results of the improved graph-based algorithm, **Fig. 8(g)** is for overlapping the foreground segmentation result on the background segmentation result together. Then the new algorithm is compared with the original algorithm [21] and Zhang, *et al*, improved graph-based algorithm [38], and the parameters are adjusted, especially *k,* to get the optimal segmentation results.

Fig. 9(a) presents the image segmentation results by the improved graph-based algorithm on the basis of guided matting algorithm. For the first image, the deer standing in the bushes, the body and the grass are mixed together, the pattern on the deer's body is perpendicular as well as the grass's texture from the vertical direction, there are differences in color barely, these are where the hard part of segmentation comes from, the proposed algorithm accurately separated the deer from the bushes; while the segmentation results of original graph based algorithm is in no order so as to cannot distinguish the foreground from background; the segmentation results of reference [38] have slight better result than that by using the original graph based algorithm, but the effect is still not satisfactory as the segmentation of the foreground from background is not accurate.
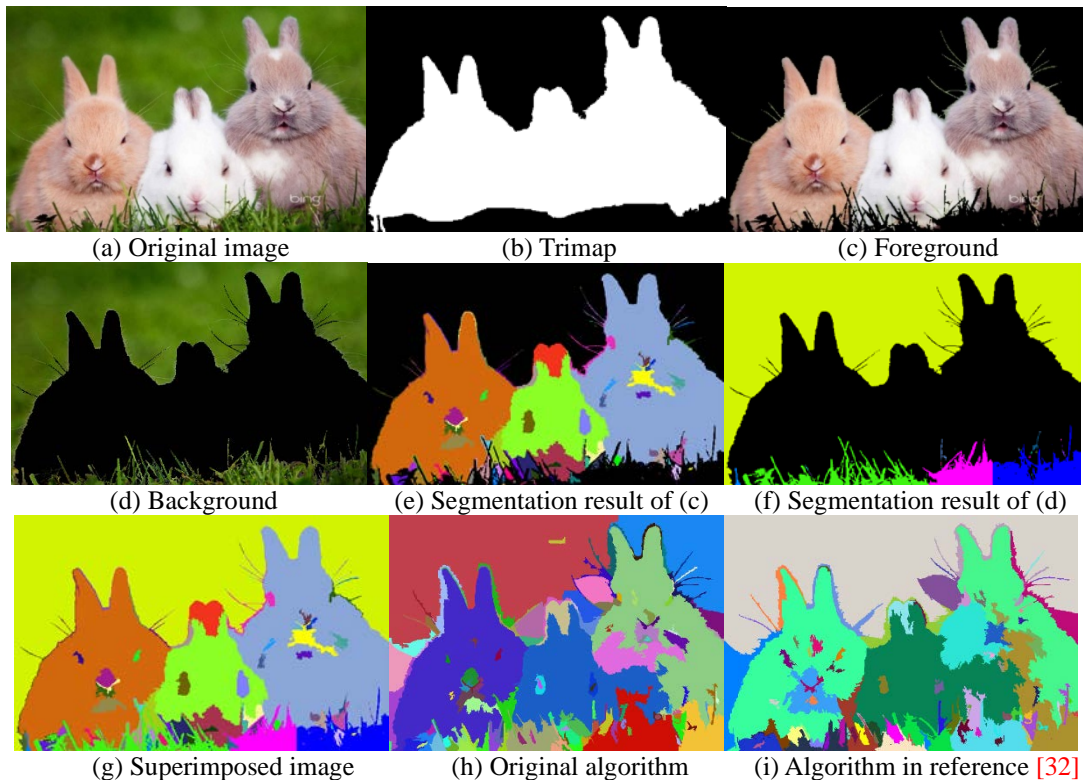

(a) Original image	(b) Trimap	(c) Foreground
(d) Background	(e) Segmentation result of (c)	(f) Segmentation result of (d)
(g) Superimposed image	(h) Original algorithm	(i) Algorithm in reference [32]

Fig. 8. Segmentation steps of the proposed algorithm and comparison to others.

In the second image, the sailboat has the similar gray scale to that in water, and it is difficult to distinguish them. Meanwhile the near leaves and far distant leaves have fuzzy boundaries that cannot be seen clearly. The difference of the gray scale values, and the gap handover between leaves and sky might cause the segmentation troubles. The proposed algorithm can overcome the above difficulties; the original graph-based algorithm produces under-segmentation problem for the foliage's gap and the clouds, the result loses a lot of details, such as several foliage at the bottom of the image; although unlike the original graph-based algorithm that causes over-segmentation problem, but the algorithm in reference [38] has serious under-segmentation problem especially for the sailing and clouds.

In the third image, the bird's body, the branches of the prospect have the resembled color with the background, it is difficult to treat, there are also small gray scale value difference between the bird's body, wings and tail feathers, as [38] not exhibit difference in bird's body and wings; the original graph-based algorithm has the same problem, there is over-merging problem in the bird body; the proposed algorithm cannot only segment the bird part exactly, but also ensure the cleanliness of the background. For the three images, according to manually drawing boundary in **Fig. 10**, the segmentation precision for each algorithm is calculated, and the comparison is shown in **Table 2**.
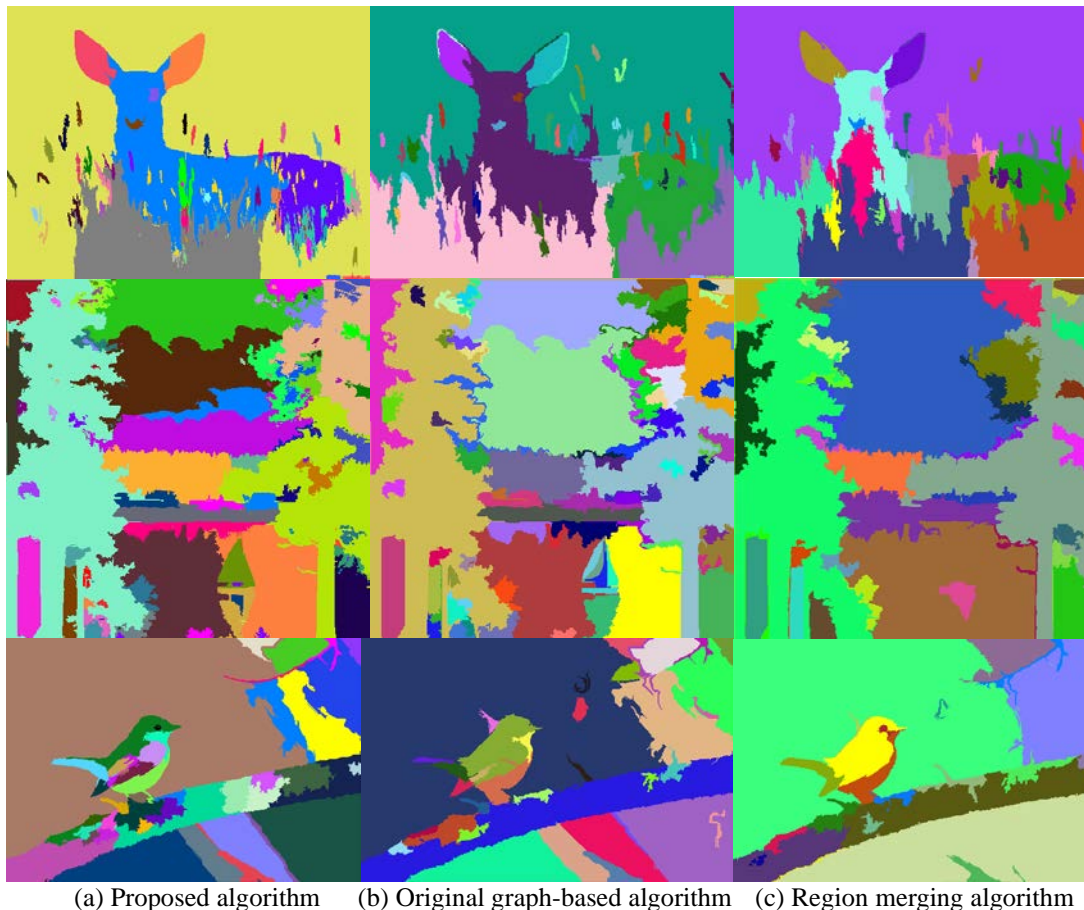

(a) Proposed algorithm   (b) Original graph-based algorithm   (c) Region merging algorithm

Fig. 9. Comparison between the proposed algorithm and others.

Accuracy equation is:

$$P = \frac{E_{pixel} - \lambda \bullet U_{pixel}}{WE_{pixel}} \qquad (20)$$

Where, E_{pixel} and U_{pixel} denote correct rates and error rates of segmented boundary points of three algorithms separately. λ is the penalty coefficient, the greater its value, the greater the punishment on the extra boundary pixels and the lower the accuracy. Taking the complexity of the image into account, all the boundary points cannot be depicted accurately, and the main consideration is for correct ratio, so this study take $\lambda = 0.3$. WE_{pixel} is the total boundary pixels of the expected segmentation results.

Table 2. Precision comparison between the proposed algorithm and others

Algorithm/Target	deer	sailboat	bird
Proposed	83.9%	87.8%	86.5%
Original	68.9%	80.3%	77.6%
Region merging	74.0%	76.2%	78.1%
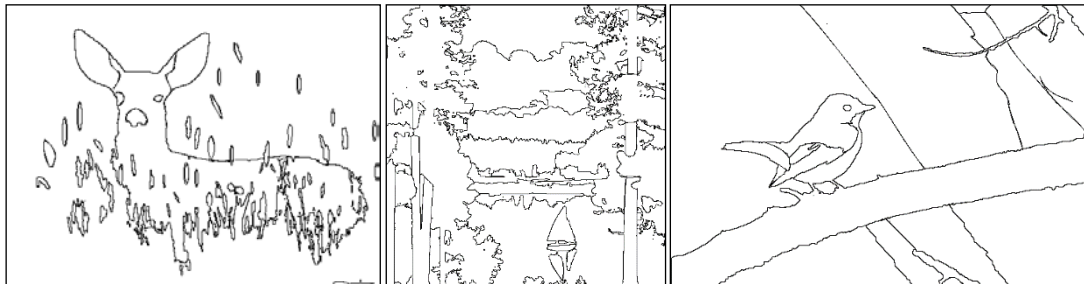
Fig. 10. Expectations of image segmentation results.

(a) Original image   (b) Proposed algorithm   (c) Original algorithm   (d) Reference [32]
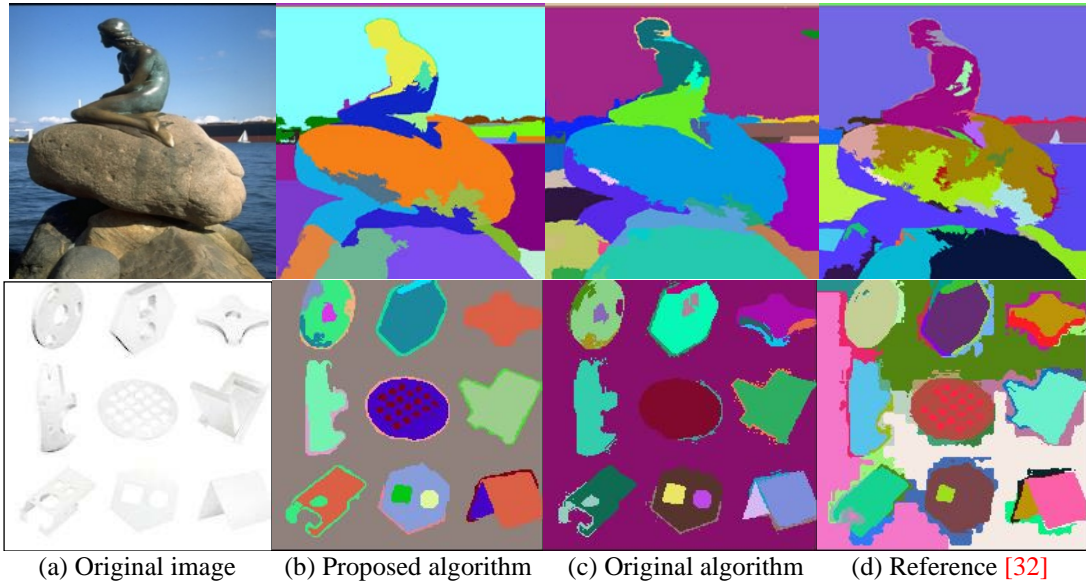
Fig. 11. Comparison of image segmentation results by three algorithms.

According to the combination of the pre-segmentation and the fine-segmentation algorithms above, the more than forty images with different characteristics are selected for tests. In this paper, the comparison of segmentation results for five images is shown in **Fig. 11**, for instance, in the bottom row, the part-type image is vague, it is difficult to distinguish the objects from background by human vision, but the image segmentation result by the new algorithm is much better than those by other algorithms. Referring to the analysis in **Figs. 9-10** and **Table 2**, obviously, the proposed algorithm has the significant advantages on the accuracy of the image segmentation.
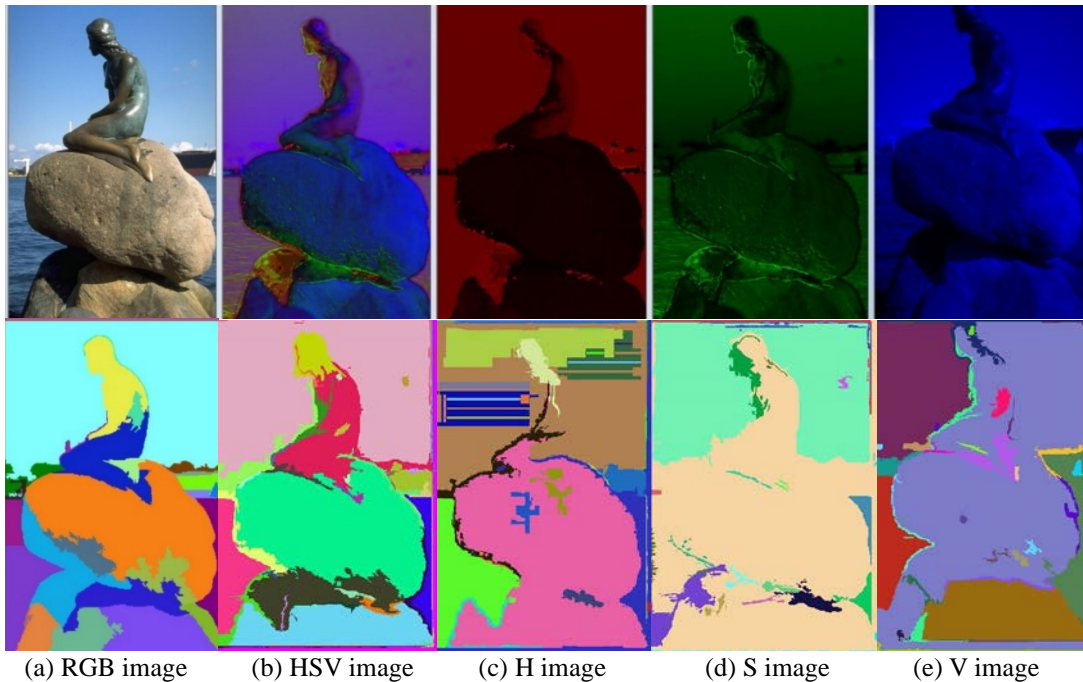

(a) RGB image   (b) HSV image   (c) H image   (d) S image   (e) V image

Fig. 12. Comparison of different images and their segmentation results: the first raw is for original images, and the second raw is for the corresponding segmentation results by the new algorithm.

In this study, for the color images, the only RGB images are used, in the future work, in order to reduce the noise, in addition to add some enhancement/smoothing algorithms or functions, the RGB color images can also be converted into the other types of images for object extraction, e.g., HSV type images, but the image converting should be image content dependent. As shown in **Fig. 12**, for the rather simple content RGB color image, the components of the HSV image have no obvious noise reducing as tested, and the segmentation results are not satisfactory as expected, even it is not good as that in the fourth row in **Fig. 11**.

5. Conclusions

The algorithm that combines the guided matting with improved graph-based image segmentation is proposed for the images where the difference between foreground and background is small. The background is extracted from foreground in an image by the guided matting sub-algorithm, then, the background image and foreground image are segmented by using the improved graph-based algorithm respectively, and finally, the two segmented images are merged together as the final segmentation result. The proposed algorithm can overall separate background from foreground correctly, and convert an image into two images. The two images remain the original image details; which is useful for the further image segmentation. Based on the detailed texture characteristics of each image, the segmentation scales and parameters can be set to correctly use the improved graph-based sub-algorithm, which can avoid under-segmentation and over-segmentation in a certain extent. Comparing to original graph-based algorithm, the improved graph-based algorithm can decrease the impact of noise and increase the image segmentation accuracy, and its main improvements are: the intra-regional and inter-regional difference functions are modified; and the edge weight function and re-merging mechanism function are added. Comparing to the ordinary region merging algorithm, the new algorithm has the lower misclassification rate and the better effects.

However, the current studied algorithm is for the images with the light vague or blurring boundaries between background and foreground, so, it needs to test more complicated images. Actually, the number of the tested images is limited (only about forty), and it requires the further testing for more types of images (e.g., hundreds to thousands). On the basis of the testing, the algorithm parameters should be automatically adjusted based on image feature analysis. In addition, the local gray scale value minima and the image fusion should also be utilized for the algorithm improvement.

Acknowledgement

References

[1] C. Zahn, "Graph-theoretical methods for detecting and describing gestalt clusters," *IEEE T. Comput.*, vol. C-20, no. 1, pp. 68-86, Jan. 1971. Article (CrossRef Link)
[2] Z. Wu, R. Leahy, "An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation," *IEEE T. Pattern Anal.*, vol. 15, no. 11, pp. 1101-1113, 1993. Article (CrossRef Link)

[3]   W. Ma, Y. Zhang, L. Yang, L. Duan, "Graph-cut based interactive image segmentation with randomized texton searching," *Computer Animation and Virtual Worlds*, vol. 27, no. 5, pp.454-465, 2016. Article (CrossRef Link)

[4]   Y. Qi Y, G. Zhang, Y. Li, "An Auto-Segmentation Algorithm for Multi-Label Image Based on Graph Cut," *Sensing and Imaging*, vol. 19, no. 1, pp.1-14, 2018. Article (CrossRef Link)

[5]   P. A. V. D. Miranda, A. X. Falcao, J. K. Udupa, "Synergistic arc-weight estimation for interactive image segmentation using graphs," *Comput Vis Image Und.*, vol. 114, no. 1, pp. 85-89, 2010. Article (CrossRef Link)

[6]   W.X. Wang, X. Zhang, T. Cao, L. P. Tian, S. Liu, Z. W. Wang, "Fuzzy and Touching Cell Extraction on Modified Graph MST and Skeleton Distance Mapping Histogram," *J. Med. Imag. Health In.*, vol. 4, no. 3, pp. 350-357(8), 2014. Article (CrossRef Link)

[7]   E.V. Dellen, I. E. Sommer, M. M. Bohlken, et al., "Minimum spanning tree analysis of the human connectome," *Hum. Brain Mapp.*, vol. 39, no. 6, pp. 2455-2471, 2018. Article (CrossRef Link)

[8]   A. Brzoza, G. Muszynski, "An approach to image segmentation based on shortest paths in graphs," in *Proc. of International Conference on Systems*, 1-5, 2017. Article (CrossRef Link)

[9]   X. Li, Y. Yang, Q. Zhao, et al., "Spatial Pyramid Based Graph Reasoning for Semantic Segmentation," in *Proc. of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, 2020. Article (CrossRef Link)

[10] J. Shi, J. Malik, "Normalized cuts and image segmentation," *IEEE T. Pattern Anal.*, vol. 22, no. 8, pp. 888-905, 2000. Article (CrossRef Link)

[11] Y. Boykov, G. Funka-Lea, "Graph cuts and efficient ND image segmentation," *Int. J. Comput. Vision*, vol. 70, no. 2, pp. 109-131, Apr. 2006. Article (CrossRef Link)

[12] E. Sharon, M. Galun, D. Sharon, et al., "Hierarchy and adaptivity in segmenting visual scenes," *Nature*, vol. 442, no. 7104, pp. 810-813, 2006. Article (CrossRef Link)

[13] X. L. Ma, L. C. Jiao, "SAR image segmentation based on watershed and spectral clustering," *J. Infrared M illim. Waves.*, vol. 27, no. 6, pp. 452-456, 2008 (in Chinese). Article (CrossRef Link)

[14] Y. K. Chen, X. M. Wu, K. Cai, et al., "CT image segmentation based in clustering and graph-cuts," *Procedia Engineering*, vol. 15, no. 1, pp. 5179-5184, 2011. Article (CrossRef Link)

[15] Y. Hou, "Research on graph theory based image segmentation," Ph.D. dissertation, School of Mechano-Electronic Engineering, XiDian Univ., Xi'an, China, 2011 (in Chinese).

[16] J. Egger, B. Freisleben, C. Nimsky, et al., "Template-cut: a pattern-based segmentation paradigm," *Sci. Rep-Uk.*, vol. 2, pp. 420, 2012. Article (CrossRef Link)

[17] W. Wieclawek, E. Pietka, "Fuzzy clustering in Intelligent scissors," *Comput. Med. Imag. Grap.*, vol. 36, no. 5, pp. 396-409, 2012. Article (CrossRef Link)

[18] V. Gopalakrishnan, Y. Hu, D. Rajan, "Random walks on graphs for salient object detection in images," *IEEE T. Image Process.*, vol. 19, no. 12, pp. 3232-3242, 2010. Article (CrossRef Link)

[19] Y. F. Yi, L. Q. Gao, L. Guo, "Mean shift based random walker interactive image segmentation algorithm," *Journal of Computer-Aided Design & Computer Graphics*, vol. 23, no. 11, pp. 1875-1881, 2011.

[20] Y. F. Wang, D. Y. Bi, F. Huang, "Improved isoperimetric algorithm for image segmentation," *Journal of XiDian University*, vol. 39, no. 2, pp. 87-94+212, 2011(in Chinese). Article (CrossRef Link)

[21] P. F. Felzenszwalb, D. P. Huttenlocher, "Efficient Graph-Based Image Segmentation," *Int. J. Comput. Vision*, vol. 59, no. 2, pp. 167-181, 2004. Article (CrossRef Link)

[22] Y. Li, H, Lu, "Natural Image Matting via Guided Contextual Attention," in *Proc. of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 07, pp.11450-11457, 2020. Article (CrossRef Link)

[23] N. Sen, A. Jain, and S. Jain, "Weighted Guided Image Filtering – A Survey," *IJCA.*, vol. 156, no. 10, pp. 29-32, Dec. 2016. Article (CrossRef Link)

[24] Y. Wexler, A. Fitzgibbon, A. Zisserman, "Bayesian Estimation of Layers from Multiple Images," in *Proc. of ECCV.*, pp. 487-501, 2002. Article (CrossRef Link)

[25] A. Levin, D. Lischinski, Y. Weiss, "A Closed-Form Solution to Natural Image Matting," *IEEE T. Pattern Anal.*, vol. 30, no. 2, pp. 228-242, 2008. Article (CrossRef Link)

[26] K. He, J. Sun, X. Tang, "Guided image filtering," *IEEE T. Pattern Anal.*, vol. 35, no. 6, pp. 1397-1409, 2013. Article (CrossRef Link)

[27] K. Hu, J. Ye, E. Fan, S. Shen, L. Huang, J. Pi, "A novel object tracking algorithm by fusing color and depth information based on single valued neutrosophic cross-entropy," *Journal of Intelligent & Fuzzy Systems,* vol. 32, no. 3, pp.1775–1786, 2016. Article (CrossRef Link)

[28] W. Wang, M. F. Wang, et al., "Pavement crack image acquisition methods and crack extraction algorithms: a review," *J. T. Transp. Eng*., vol. 6, no. 6, pp. 535-556, 2019. Article (CrossRef Link)

[29] W. Wang, H. X. Li, et al., "Pavement crack detection on geodesic shadow removal with local oriented filter on LOF and improved Level set," *Constr. Build. Mater.*, vol. 237, no. 2020, pp. 117750, 2020. Article (CrossRef Link)

[30] W. X. Wang, W. W. Chen, et al., "Extraction of tunnel centerline and cross sections on Fractional calculus and 3D invariant moments and best-fit ellipse," *Opt. Laser Technol.*, vol. 128, no. 2020, pp. 106220, 2020. Article (CrossRef Link)

[31] G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," *IEEE T. Pattern Anal.*, vol. 40, no. 4, pp. 834-848, April. 2018. Article (CrossRef Link)

[32] S. Liu, M. Rahman, C. Lin, C. Wong, G. Jiang, S. Liu, N. Kwok, H. Shi, "Image contrast enhancement based on intensity expansion-compression," *Journal of Visual Communication and Image Representation*, vol. 48, pp. 169-181, 2017. Article (CrossRef Link)

[33] Z. He, Y. Cao, L. Du, B. Xu, J. Yang, Y. Cao, S. Tang and Y. Zhuang, "MRFN: Multi-Receptive-Field Network for Fast and Accurate Single Image Super-Resolution," *IEEE Transactions on Multimedia*, vol. 22, no. 4, pp. 1042–1054, 2020. Article (CrossRef Link)

[34] J. Pi J, K. Hu, Y. Gu, L. Qu, F. Li, X. Zhan, Y. Zhan, "Robust scale adaptive and real-time visual tracking with correlation filters," *IEICE TRANSACTIONS on Information and Systems*, vol. 99, no. 7, pp.1895-902, 2016. Article (CrossRef Link)

[35] Q. Li, Z. Cao, W. Ding W, et al., "A multi-objective adaptive evolutionary algorithm to extract communities in networks," *Swarm and Evolutionary Computation*, vol. 52, 100629, 2020. Article (CrossRef Link)

[36] L. Fan, E. Fan, C. Yuan, K. Hu, "Weighted fuzzy track association method based on Dempster-Shafer theory in distributed sensor networks," *International Journal of Distributed Sensor Networks*, vol. 12, no. 7, 10 Pages, 2016. Article (CrossRef Link)

[37] Q. Li, Z. Cao, J. Zhong and Q. Li, "Graph representation learning with encoding edges," *Neurocomputing*, vol. 361, pp. 29–39, 2019. Article (CrossRef Link)

[38] X. Long, J. Sun, "Image segmentation based on the minimum spanning tree with a novel weight," *Optik - International Journal for Light and Electron Optics*, vol. 221, 165308, 2020. Article (CrossRef Link)

Weixing Wang is the professor in information engineering. He obtained his PhD degree in 1997 at Royal Institute of Technology in Sweden, where, since 2001 until now, he has been a PhD supervisor. Now he is also a visiting professor at Shaoxing University. His research interests include information engineering, image processing and analysis, pattern recognition and computer vision.

Angyan Tu received the M.S. degree in computer science and technology from Zhejiang University, and now as a doctoral student is in Shanghai University. She is currently a lecturer at Shaoxing University. Her current research interests include image processing and analysis, neutrosophic theory, scheduling and optimization algorithms.

Fredrik Bergholm is the professor in computer science and technology. He obtained his PhD degree in 1987 at Stockholm University in Sweden. Since 1991 until now, he has been a professor and a PhD supervisor at Royal Institute of Technology. His research interests include image processing and analysis, pattern recognition, computer vision, numerical analysis and computer simulation.