Multi-scale Gaussian Segmentation via Graph Cuts

Ye Zhang* and Kun He
College of Software, Jiang’an campus of Sichuan University, 610225 Chengdu, China

ABSTRACT

Towards the improvement of image segmentation performance, a new improved segmentation method is proposed which combines the use of multi-scale information. The multi-scale information of an image, generated by a linear Gaussian smoothing function through an interactive scheme that provides different levels of image cues which help to capture the abundant content of the image. However, Gaussian smoothing cannot generate an optimum solution to the specific scale level. According to the boundaries sensibility to scale, a convergence constraint based on the significant scale level of the segmented regions is constructed. The experimental results show that this method has superior advantages in performance and towards noise images.

INTRODUCTION

Image segmentation is a process of subdividing an image into several sub-regions based on contents. Effective segmentation technologies provide image retrieval and object analyzing with useful information [1]. They have been widely used in fields like medical analysis [2] and object localization [3].

There are many segmentation technologies based on graph theory, such as N-cuts [4], Grab cuts [5] and Graph Cuts [6] [7] [8] [9]. The segmented regions using N-cuts have no semantic concept for only exploiting low-level nature information. With the help of user interaction [10], Grab Cuts and Graph Cuts extract semantic objects combining boundary and regional information. In Grab Cuts, regional energy function has locality because of regional Gaussian fitting. It leads to local optimum of the segmentation results. Graph Cuts uses histograms to obtain global information, which improves the segmentation performance in comparison with Grab Cuts. However, both cannot produce correct segmentation results on noisy images or images with overlap appearance between objects and background owing to taking given-scale information into account [11]. The full convolution nets [12] which exploits different-scale information, is proposed to implement semantic image segmentation. However, it needs massive image data which is accurate to pixel level [13] for deep learning.

Inspired by the Graph Cuts, we propose an improved segmentation method. In this method, the multi-scale information [14] is generated by Linear Gaussian smoothing, which causes component with more peaks. It is benefit for segmentation performance. However, Gaussian smoothing is linear, leading there is not a solution to the optimum scale problem. According to boundary sensibility to scale, a convergence condition on the significant level of segmentation regions is constructed. The experimental results show that the proposed method has superior performance and is insensitive to noise.
RELATED WORK ON MULTI-SCALE INFORMATION

Image segmentation exploiting single-scale information cannot produce correct results. In order to improve the segmentation performance, we use Gaussian smoothing to obtain multi-scale information. Let \( I^0 \) denotes the luminance component of a color image and \( G_\sigma \) is the Gaussian function [15]. Smoothing component \( I \) is given by

\[
I = I^0 * G_\sigma
\]  

(1)

where \( \sigma \) is the standard deviation of Gaussian and decides the smoothing scales. With \( \sigma \) increasing, the small-scale objects of an image are smeared to large structures step by step. It is difficult to get an appropriate \( \sigma \) without prior knowledge. To optimize \( \sigma \), we use iteration to produce a series of components \( I^k \) using Gaussian kernel \( G_\sigma \ k = 0,1,\ldots,k,\ldots \) according to equation 1. To reduce computation cost, according to the Gaussian sequential convolution, that is \( \sigma^2 = \sigma_0^2 + \Delta \sigma^2 \), the different-scale smoothing component \( I^k \) can be obtained based on the last smoothing component \( I^{k-1} \), instead of original image \( I^0 \), just as follows:

\[
I^0 \xrightarrow{\Delta \sigma} I^1 \xrightarrow{\Delta \sigma} \cdots \xrightarrow{\Delta \sigma} I^k \xrightarrow{\Delta \sigma} \cdots
\]  

(2)

where \( \sigma_0 \) is the original standard deviation of Gaussian applied to original image \( I^0 \) and \( \Delta \sigma \) is applied directly to the last smoothing component.

At step \( k \), smoothing component \( I^k \) only depends on \( I^{k-1} \), leading that the small-scale objects are smeared step by step, the appearance of \( I^k \) is more tight cohesion than that of \( I^{k-1} \), the histograms of components appear peak distribution (the second column in Figure 1). If smoothing scale increases continuously, the distribution of histogram is almost invariant but edges are blurred (the bottom-left corner sub-image in Figure 1).

PROPOSED SEGMENTATION MODEL

Graph Cuts can get global optimum for image segmentation. We exploit it to segment the smoothing component \( I^k \). The segmentation can be estimated as a global minimum by standard Minimum Cut algorithm [16]:

\[
E(I^k) = \lambda R(I^k) + B(I^k)
\]  

(3)

where \( R(I^k) \) donates regional information, which is

\[
R(I^k) = \sum_{i=1}^{N} R_p(I^k_i)
\]  

(4)

Here \( R_p(I^k_i) \) represents the individual penalties for assigning pixel \( i \) to "object" or "background". \( B(I^k) \), which attracts the segmentation curve to edge, is set by a decreasing function to switch appropriately between high and low contrast, which is
\[ B(I^k) = \sum_{(p,q) \in G} \frac{\gamma}{\text{dist}(p,q)} \times \exp(-\eta \times \| (I^k_p - I^k_q) \|^2) \]  

where \( G \) represents the set of all the neighboring pixels pairs and \( \text{dist}(p,q) \) is the Euclidean distance between pixels \( p \) and \( q \). The \( \eta \) should be calculated to switch appropriately between high and low contrast edge, that is

\[ \eta = 0.5 \times \left\langle (I^k_p - I^k_q)^2 \right\rangle^{-1} \]

where term \( \left\langle \cdot \right\rangle \) is expectation over the smoothing component \( I^k \).

The multi-scale information is obtained by linear Gaussian kernel, leading that there is not a solution to the scale optimum problem. An image from the international web was smoothed with different number of iterations. Part results are shown in Figure 1. With smoothing scale increasing, performance in high-scale segmentation is superior to that in low-scale segmentation (seen the red oval in the first and second row). When the number of iteration keeps on increasing, improper-scale information causes decrease of the segmentation performance (seen the blue oval in the third row).

In order to evaluate segmentation performance, we construct a convergence condition based on the significant level of segmentation sub-regions on adjacent scale components. It is defined as follows:

\[ C(S^k, S^{k-1}) = \frac{\text{card}(S^k \cap S^{k-1})}{\max\{\text{card}(S^k), \text{card}(S^{k-1})\}} \]

where \( C(S^k, S^{k-1}) \) represents the significant level of segmentation performance on two adjacent scale component \( I^k \) and \( I^{k-1} \). \( S^k \) represents the set of pixels which belong to object regions in the segmentation result when segmenting the smoothing component \( I^k \). \( \text{card}(S^k) \) represents the number of elements in the set \( S^k \). If the difference of
segmentation performance between two adjacent scale components decrease, $C(S^k, S^{k-1})$ will increase. With scale increasing, the appearance has more peak distribution, leading that the significant level of the segmentation increases monotonically. If the iteration number tends to infinitely great, regional diversity disappears, which causes the significant level to monotonously decrease. When the confidence level satisfies the following condition, the smooth scale is terminated (see Figure 2):

$$C(S^k, S^{k-1}) - C(S^{k-1}, S^{k-2}) \leq 0$$  \hspace{1cm} (8)

Algorithm1: Procedural steps of the proposed method

| Initialized parameters - $I^0, \sigma_0, \Delta\sigma, S^0, \gamma, \lambda$ |
|---|---|
| Begin |
| Selecting region of interest and background regions and Initializing $S^0$ |
| Do for each $k$ scale [do … while loop] |
| Applying Gaussian smoothing through the equation 2 |
| Substitute in Gaussian component $I^k$ into equation 3 |
| While (Until convergence is reached) given by equation 8 |
| Save output results and computation times. |
| End |

Figure 2. Procedural steps of the proposed method.

EXPERIMENTAL RESULTS

The experiments were conducted according to Algorithm1 using C language in VC6.0 on the PC with Duo Core i5 CPU @2.0GHZ and 4GB of RAM. During the experiment, $\gamma$ in equation 5 was set to 50 and since the edge boundary and region information is equally important, $\lambda$ was set to 4.0 in equation 3. Images were segmented using the Graph Cuts on multi-scale smoothing components and the parameters $\sigma_0$ and $\Delta\sigma$ decide the smoothing scale. To analyze the relationship between these two parameters and the segmentation performance, an image in Berkeley segmentation database (https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/BSDS300/html/dataset/images/color/test-026-050.html) was smoothed with different $\sigma_0$ and $\Delta\sigma$, and partial results are shown in Figure 3. Ground truth represents the correct segmentation result. As we can see, with smoothing scale increasing, segmentation performance is getting better first and then getting worse. In the following experiments, $\sigma_0$ and $\Delta\sigma$ were set to 0.6 by optimizing performance against ground truth over 15 images.

In order to test the segmentation performance using proposed method on real images, experiments were carried on to compare with Graph Cuts model. Partial results are shown in Figure 4. The CPU time and score of segmentation (precision, recall and F-Measure) using these two methods are given in Table I. Precision is the number of returned pixelated information that are correct and whereas recall is the number of positively correct pixelated information returned by that model. F-Measure is the equation between these two evaluating indicators. The bigger F-Measure is, the better the segmentation performance is. As we can see, for images with strong contrast between objects and background (like the first row in Figure 4), the segmentation
performance using proposed method is almost the same as that using Graph Cuts model. The effect of the proposed method on images in which there is a wide range of luminance in object region, is better than that of the Graph Cuts (the red oval in row 2). For images with overlap appearance between objects and background, in comparison with Graph Cuts, the proposed method is out-performance (the last row) because the appearance of the multi-scale component have more peak distributions. However, as we can see in TABLE I, the proposed method is computationally longer than the Graph Cuts.

Noise is one of the factors causing poor segmentation performance. In order to test robustness against noise using these two methods, we conduct experiments on images with additive Gaussian noise. Partial results are shown in Figure 5. The scores of segmentation on images with and without noise are listed in TABLE II. For original image, the performance of both methods is almost the same. The appearance distribution of the noisy image causes the negative effects on segmentation performance using Graph Cuts model. The segmentation results using the proposed method are out-performance compared to those using the Graph Cuts because multi-scale smoothing ameliorate the appearance distribution of the noisy images (the fourth and sixth column). Although the proposed method is insensitive to noise, the CPU time is longer than Graph Cuts. The lower the PSNR, the longer the CPU time.

![Figure 3. Segmentation results with different parameters $\sigma_0$ and $\Delta \sigma$.](image)

![Figure 4. Comparison of the proposed method with Graph Cuts.](image)
## TABLE I. SCORE OF SEGMENTATION AND EXECUTE TIME COMPARISON OF FIGURE 4.

| Methods      | Images     | The Graph Cuts model | The proposed method |
|--------------|------------|-----------------------|---------------------|
|              |            | precision | recall | F-Measure | CPU | precision | recall | F-Measure | CPU |
| flower       | 512×384    | 0.999     | 0.998  | 0.993     | 5.19 | 0.999     | 0.986  | 0.992     | 9.87 |
| panther      | 192×229    | 0.999     | 0.862  | 0.926     | 1.13 | 0.992     | 0.948  | 0.969     | 4.19 |
| squirrel     | 340×191    | 0.750     | 0.924  | 0.828     | 2.62 | 0.952     | 0.930  | 0.941     | 7.66 |

Figure 5. Comparison of the proposed method with Graph Cuts on real images with noise. Column 1-6, user interaction, ground truth, the proposed method, the histograms of the smoothing component, the Graph Cuts model, and the histograms of the image. Row 1-3: original image, PSNR (Peak Signal Noise Ratio) = 21.26, PSNR = 14.01.

## TABLE II. SCORE OF SEGMENTATION AND CPU TIME COMPARISON OF NOISY IMAGE.

| PSNR  | The Graph cuts model | The proposed method |
|-------|----------------------|---------------------|
|       | precision | recall | F-Measure | CPU | precision | recall | F-Measure | CPU |
| original image | 0.996 | 0.999 | **0.993** | 3.02 | 0.995 | 0.947 | 0.970 | 8.44 |
| 26.53 | 0.998 | 0.974 | **0.981** | 3.88 | 0.987 | 0.901 | 0.942 | 8.26 |
| 23.33 | 0.988 | 0.907 | 0.946 | 5.25 | 0.994 | 0.903 | **0.958** | 12.58 |
| 23.08 | 0.976 | 0.909 | 0.942 | 6.42 | 0.996 | 0.922 | **0.958** | 13.18 |
| 21.26 | 0.988 | 0.908 | 0.946 | 9.26 | 0.985 | 0.928 | **0.956** | 15.60 |
| 18.46 | 0.996 | 0.714 | 0.832 | 11.60 | 0.993 | 0.908 | **0.920** | 20.51 |
| 14.01 | 0.981 | 0.611 | 0.753 | 16.13 | 0.975 | 0.927 | **0.958** | 23.37 |
| average | 0.989 | 0.860 | 0.913 | 7.93 | 0.989 | 0.919 | **0.950** | 14.56 |

## CONCLUSIONS

The use of combined multi-scale information in the proposed method improved the intensity histogram of an image to a point that an optimum solution could be obtained. This gives a distinctive peak distribution for image segmentation which leads to superior performance even under noisy conditions. This proposed method could be well suited for low resolution camera applications such as in Robotics. However, there are some drawbacks in the proposed method. 1) The proposed method is computationally longer than the current method. 2) The Gaussian kernel is isotropy diffusion, leading that the segmentation curve positions the original contour imprecisely, and the segmentation results on images with the weak edge are unsatisfactory.
ACKNOWLEDGEMENT

This work was supported by the Sichuan Province Natural Science Foundation of China (2016JZ0014).

REFERENCES

1. Zhang X., Su H., and Yang L., et al. 2015. “Fine-grained histopathological image analysis via robust segmentation and large-scale retrieval,” Computer Vision and Pattern Recognition IEEE, 29(5):5361-5368.
2. Masood S., Sharif M., and Masood A., et al. 2015. “A survey on medical image segmentation,” Current Medical Imaging Reviews, 11(1): 3-14.
3. Brejl M. and Sonka M. 2000. “Object localization and border detection criteria design in edge-based image segmentation: automated learning from examples,” IEEE Transactions on Medical Imaging, 19(10): 973-985.
4. Shi J. and Malik J. 2000. “Normalized cuts and image segmentation,” IEEE Computer Society, 22(8):888-905.
5. Rother C., Kolmogorov V. and Blake A. 2004. “‘GrabCut’: interactive foreground extraction using iterated graph cuts,” ACM SIGGRAPH, 23(3):309-314.
6. Boykov, Y. Y. and M. P. Jolly. 2001. “Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images,” Proc. Int’l Conf. Computer Vision, 1, pp. 105-112.
7. Wang T., Ji Z., and Sun Q., et al. 2016. “Label propagation and higher-order constraint-based segmentation of fluid-associated regions in retinal SD-OCT images,” Information Sciences, s 358-259(C), 92-111.
8. Wang T., Sun Q., and Ji Z., et al. 2016. “Multi-layer graph constraints for interactive image segmentation via game theory,” Pattern Recognition, 55(C):28-44.
9. Wang T., Ji Z., and Sun Q., et al. 2016. “Interactive Multilabel Image Segmentation via Robust Multilayer Graph Constraints,” IEEE Transactions on Multimedia, 18(12): 2358-2371.
10. Ning J., Zhang L., Zhang D. and Wu C. 2010. “Interactive image segmentation by maximal similarity based region merging,” Pattern Recognition, 43(2):445–456.
11. Shi J. and Malik J. 2002. “Normalized cuts and image segmentation,” IEEE Computer Society, 22(8):888-905.
12. Kalinovsky A. and V. Kovalev. 2016. “Lung Image Segmentation Using Deep Learning Methods and Convolutional Neural Networks,” presented at Xiii International Conference on Pattern Recognition and Information Processing.
13. Long J., Shelhamer E., and Darrell T. 2015. “Fully convolutional networks for semantic segmentation,” Computer Vision and Pattern Recognition IEEE, 79(10):3431-3440.
14. Zheng S., Yuille A. and Tu Z. 2010. “Detecting object boundaries using low-, mid-, and high-level information,” IEEE Conference on Computer Vision and Pattern Recognition, 114(10):1-8.
15. Witkin A. P., Baudin M. and R. O. Duda. 1986. “Uniqueness of the Gaussian kernel for scale-space filtering,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(1): 26-33.
16. Boykov Y. and Kolmogorov Y. 2004. “An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(9):1124-1137.