ROBUST INTERACTIVE IMAGE SEGMENTATION VIA ITERATIVE REFINEMENT

Yao Peng[1], Juyong Zhang[1], Yancheng Yuan[1], Shuyuan Zhu[2], Lu Fang[1]

[1]USTC, [2]UESTC

ABSTRACT

Image segmentation with user inputs gets more and more popular in recent years and always performs better compared with automatic methods. However, existing interactive image segmentation methods still might fail if the image contains messy textures, or the user inputs are sparse or at inappropriate locations. In this paper, we propose a novel iterative refinement framework which leads to robust segmentation performance even with sparse and improper input strokes. Specifically, a geodesic distance based energy is introduced and combined with convex active contour model, and an iterative seeds refinement technique is put forward to handle the sparse input problem. Extensive experiments using real world images, and segmentation benchmark dataset show that our proposed method has superior performance compared with representative state-of-the-art methods.

1. INTRODUCTION

1.1. Motivation

In recent years, interactive image segmentation becomes more and more popular as it can produce more reliable segmentation result with the help of the information supplied by users. This type of interactive way is quite useful for many image editting problems and medical image processing applications. However, the performance of interactive image segmentation heavily depends on the quality and quantity of the inputs seeds supplied by the users. For most of existing methods, the regions without seeds might be assigned to incorrect labels. For an object with complex texture, the input seeds can not cover every sub-region of the interested object. These regions without seeds can not be segmented robustly. Moreover, the seeds are fixed during the whole segmentation process in most of existing methods. In other words, much information of former iterations are not involved in the following iterations and these regions without seeds can not be segmented robustly. In this paper, we try to propose a robust interactive image segmentation method to deal with these problems which are rarely considered before.

1.2. Related Works

In general, there are two categories of interactive image segmentation approaches: boundary-based and region-based ones. State-of-the-art region-based interactive segmentation algorithms include Graph Cut based methods [1, 2], Random Walks(RW) based methods [3, 4], and Geodesic methods [5, 6]. All these methods firstly convert an image into a weighted graph and then minimize an objective energy functional to produce the final segmentation. However, the Graph Cut algorithm is sensitive to the number of seeds, while the RW and Geodesic algorithms are sensitive to the locations of seeds [7].

In 2007, Bresson et. al. [8] proposed a convex active contour model based on "active contours without edges" model [9] and "Mumford-Shah" model [10], utilizing both boundary and regional information. Recently, Nguyen et. al. [11] extended it to interactive image segmentation, which is named constrained active contour model. Specifically, the segmentation result of Geodesic method [5] is used for contour initialization and foreground/background Gaussian Mixture Models(GMMs) [2] are utilized to represent regional term.
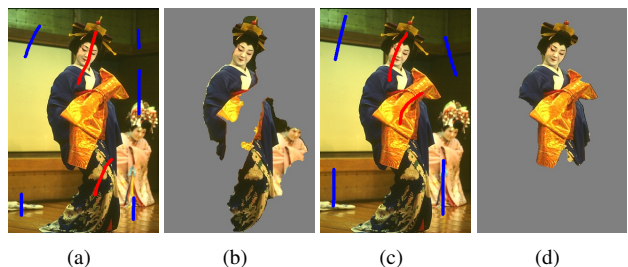

(a)   (b)   (c)   (d)

Fig. 1. An illustration of different segmentary results using [11], under different sparse input strokes: (a)(c) different input seeds; (b)(d) segmentary results of input seeds in (a)(c) respectively.

The energy term in segmentation models is always composed by global energy term [12, 13] and local energy term [14, 15]. The former one relates each pixel to all samples and it can quickly propagate seeds through the image while the latter one involves more user interaction. Recently a new method proposed by Xu [16] is based on iterative feature discrimination and it relates each pixel to part of the control samples. It successfully combine the effect of both global and local operators even with a small amount of user inputs.

Through extensive experiments we find the crucial problems of existing methods are: 1) the results of segmentation are extremely sensitive to the locations and quantity of seeds;

2) these methods can barely work on images with complex color distribution or messy textures. Results of using a representative method [11] with different sparse user inputs are shown in Fig. 1. Obviously, the different performances in Fig. 1(b) and (d) indicate that the segmentation is sensitive to locations of seeds, especially without sufficient input strokes.

1.3. Our Contributions

Current interactive image segmentation methods mainly consider the color information like the tensity or their distributions, and the user supplied seeds are fixed and used only once. The fixed user inputs might result in uncertainties when the desired segmentation region does not contains input seeds. Besides, as mentioned in the motivation part, invariable seeds will lose some important information during the iterations in the segmentation procedure. In this paper, a new model with seeds refinement is proposed and performs much better, even for images with sparse inputs. In summary, the main contributions of this paper are:

- A geodesic energy term is introduced and combined with convex active contour model to enhance the weightiness of those seed regions.

- A seeds refinement technique is proposed to enhance the reliable segmentation results, and use them to guild the unreliable part segmentation.

- Experiments on MSRC dataset show that the proposed method produces high quality segmentation results even with few user inputs.

The rest of paper is organised as follows: in Section 2 we explain new Geodesic energy term together with constrained active contour model as well as the proposed iterative seeds refinement scheme. In Section 3, we verify our method by showing some representative results and discussions. The conclusions are given in Section 4.

2. ROBUST INTERACTIVE IMAGE SEGMENTATION

In this section, we will first introduce how to modify the Constrained Active Contour Model to incorporate the seeds location information. Then the interative seeds refinement strategy is proposed to handle the low quality and small quantity of input seeds problem.

2.1. Modified Constrained Active Contour Model

Active contour model [11, 17] gets big success in image segmentation. However the seed locations are not emphasized in existing models, and thus the regions without seeds might be assigned with incorrect labels. To handle this problem, a geodesic energy term is added to enhance the weightiness of those seed regions. The proposed model is formulated as

$$\min_{u\in[0,1]} \quad \int_\Omega g_b|\nabla u| + \lambda_1 h_r u + \lambda_2(d_F - d_B)u^2 dx$$
$$\text{s.t} \quad u(x) = 1 \quad \text{if } x \in F \tag{1}$$
$$u(x) = 0 \quad \text{if } x \in B,$$

where λ_1 and λ_2 are two trade-off parameters. u is a probability function defined on image domain Ω, which has a value between 0 and 1 at each pixel location x in the image. Set F(or B) is the pixels in foreground(or background). The segmented region is obtained by thresholding the function u. Different from the models in [11, 17], the geodesic energy term $(d_F - d_B)u^2$ is added to the modified constrained active contour model to enhance the performance, where $d_F(x)$ and $d_B(x)$ are the geodesic distance of pixel x to foreground and background [5] separately. g_b and h_r in our model are the same as in [11], which describe the boundary and regional properties of the image,

$$g_b = \beta \cdot g_c + (1 - \beta) \cdot g_e,$$
$$h_r(x) = \alpha(P_B(c_x) - P_F(c_x)) + (1 - \alpha)(1 - 2P(x)),$$

where g_c and g_e are edge detection results of $P_F(c_x)$ and the original image, respectively. $Pr(c_x|F)$, $Pr(c_x|B)$ denote the probabilities of the Gaussian Mixture Models(GMMs). The likelihood of a pixel x with color c_x belonging to the foreground becomes $P_F(c_x) = Pr(c_x|F)/(Pr(c_x|F) + Pr(c_x|B))$. $P(x) = D_B(x)/(D_F(x) + D_B(x))$ is assigned to $u(x)$ as initialization.
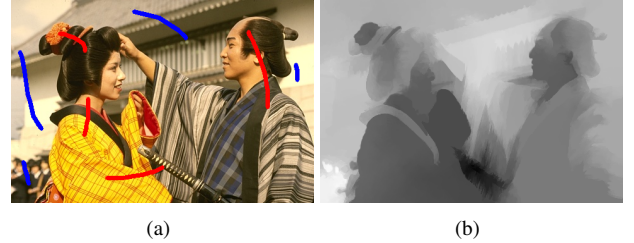

(a)   (b)

Fig. 2. (a) A sample image with sparse inputs; (b) the visualization of $d_F - d_B$.

The newly added geodesic energy term in model Eq. 1 increases the weight of the region near the seeds. If $d_F(x) < d_B(x)$, it means that pixel x is closer to the foreground seeds from the perspective of geodesic distance, and thus $u(x)$ is supposed to assign a larger value. This term magnify the effect of the seed regions. As shown in Fig. 2, the visualization of $d_F - d_B$ is showed. The week edges such as the texture does not increase the geodesic distance. Thus the region without seeds but near with the seeds region in the geodesic distance can obtain a more proper segmentation result.

To solve Eq. 1, we first introduce an auxiliary variable d to substitute ∇u and rewrite it as

$$\min_{u\in[0,1]} \quad \int_\Omega g_b|d| + \lambda_1 h_r u + \lambda_2(d_F - d_B)u^2 dx$$
$$\text{s.t} \quad d = \nabla u; \tag{2}$$
$$u(x) = 1, \quad \text{if } x \in F;$$
$$u(x) = 0, \quad \text{if } x \in B.$$

Eq. 2 can be optimized via Split Bregman [18], i.e.,

$$(u^{k+1}, d^{k+1}) = arg \min_{u \in [0,1], d} \int_{\Omega} g_b |d| + \lambda_1 h_r u +$$
$$\lambda_2 (d_F - d_B) u^2 + \frac{\mu}{2} |d^k - \nabla u - b^k|^2 dx, \quad (3)$$

$$b^{k+1} = b^k + \nabla u^{k+1} - d_{k+1}, \quad (4)$$

where superscript k indicates the iteration index, and b is the dual variable. Based on Euler-Lagrange differential equation, the value of u can be solved by a Gauss-Seidel iterative method,

$$\mu \nabla u - 2\lambda_2 (d_F - d_B) u = \lambda_1 h_r + \mu \text{div}(d^k - b^k), \quad (5)$$

where div is the divergence operator. The minimum solution of d value in Eq. 3 can be obtained by soft-thresholding,

$$d^{k+1} = \frac{\nabla u^{k+1} + b^k}{|\nabla u^{k+1} + b^k|} \max(|\nabla u^{k+1} + b^k| - \mu^{-1} g_b, 0). \quad (6)$$
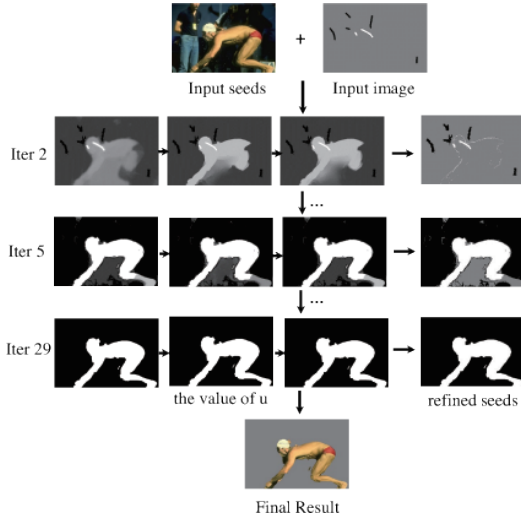
2.2. Iterative Seeds Refinement

Fig. 3. Pipeline of our method. The white pixels represent foreground seeds and the black ones represent background. Within each iteration, u is gradually refined as shown in each row, and the seeds are updated based on the optimized u.

In existing interactive image segmentation methods, the final segmentation result is decided by optimizing the proposed energy once with the initial user inputs. However, when the user input strokes are sparse or its locations are inappropriate, the complex objects are hardly to be segmented correctly. We propose to apply model Eq. 1 to calculate the $u(x)$ multiple times by iteratively refining the seeds. With the newly updated seeds based on the optimized $u(x)$ value, we update the foreground/backgroud GMMs and recalculate the geodesic distance for every pixel to the updated seeds in each iteration. And then, the probability function $u(x)$ is recomputed by Eq. 1. This iteration is continued until the proportion of seed regions exceeds the predefined threshold. The final segmentaiton can be got through applying the threshold T on the final $u(x)$. The seeds refinement and model's updating process are illustrated in Fig. 3 and our demo [19].

Instead of segmenting images based on the initial input seeds in the model whose energy function is minimized only once, we make full use of the information which can generate the new seed regions properly and then the active contour model is updated. The more seeds input to the model, the more regions can be determined with high assurance. We iteratively update the segmentation results by solving the model multiple times with refined seeds. Fig. 4 indicates obviously that with more iterations, the better visual results are achieved.
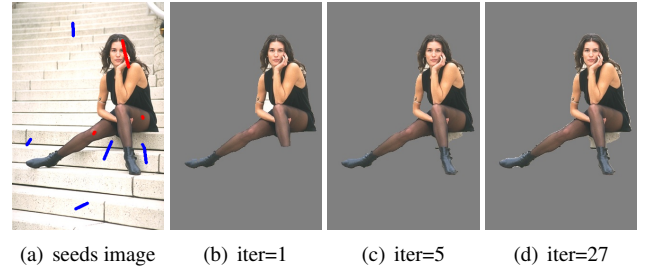

(a) seeds image   (b) iter=1   (c) iter=5   (d) iter=27

Fig. 4. Results at different iterations for image with complex color distribution and texture.

3. EXPERIMENTAL RESULTS AND DISCUSSIONS

Similar to the settings in [11], λ_1 is set to 100 in our experiments. Different λ_2 values are evaluated as well. For over small λ_2, the geodesic energy does not take important effect. While large λ_2 leads over dominant of seed locations. In the following experiments, we set λ_2 to 1000. The commonly used MSRC ground truth data set [2] that contains 50 test images is used for comparison with existing methods.

Fig. 5 describes our segmentation results for image with complex foreground and background, under sparse inputs on image. Note that the sparse inputs indicates that we merely choose at most 1% pixels of the whole image. Apparently, the results with our proposed model gets better with more iterations, while the results generated by [11] are unsatisfied, due to complex image contents and limit quantity of the input seeds. With the geodesic energy term, the texture structure around with foreground seeds can benefit regions which are within local area of seeds. For cases with similar foreground and background, the method in [11] can hardly find the exact cutting contour, or even completely fail in some particular situations by merely segmenting stroke regions, such as the third example in Fig. 5. With the help of more iterations in our proposed method, the segmentation has a chance to correct error, as shown in the last column of Fig. 5.
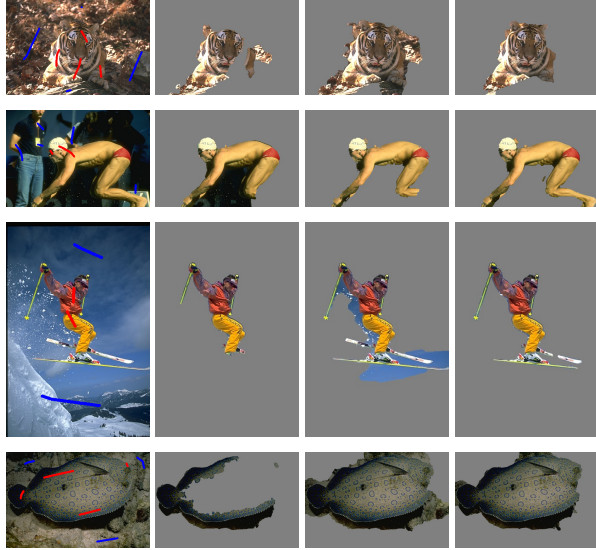
Fig. 5. Segmentation results on complex images with sparse strokes. The second column are the results of [11]. The third column shows the results of model in Eq. 1. The last column shows the final results until seed refinement stops.

In addition, we testify our method under different seed locations as shown in Fig. 6. Although the seeds do not cover every component of foreground, and the seeds are quite different in both cases, both of the two results are similar and satisfying, indicating that our method is more robust to sparse input strokes than some state-of-art method (such as the one in Fig. 1).

Fig. 6. With different inputs (same with Fig 1), the segmentation results of proposed method are satisfying and similar.

Table 1 reports the error rates (percentage of mislabeled pixels) of concerned methods and our proposed method. The error rate is computed via the average result of 50 images in MSRC dataset. We test the Constrained Active Contour Model (Constrained ACM) [11] with both normal and sparse inputs for comparison. It is obvious that the proposed method achieves lower error rates than Constrained ACM for both sparse and normal inputs. All the input sparse strokes and the segmentation results are supplied in [20]. Note that the performances of Grabcut [2], RW with AT [21] and Constrained ACM [11] are under exactly the same conditions as where they are reported.

Method	error rate (%)
Grabcut [2]	5.66 (reported in [11])
RW with AT [21]	3.3 (reported in [21])
Constrained ACM [11]	3.77 (reported in [11])
Constrained ACM [11]	5.06 (with sparse inputs)
Our Method	**4.30** (with sparse inputs)
Our Method	**2.78** (with normal inputs)

Table 1. Error rate comparison using the MSRC dataset.

Although the iterative seeds refinement and optimization of Eq. 1 are required in the proposed method, the numerical algorithm in Eq. (6) is quite efficient, and the value of u from previous iteration serves as a good initial value. The corresponding computational efficiency for representative images (in Fig. 5) is listed in Table2, under the computer setting of Inter Core i7-3770K CPU and 32GB RAM.

Image	Size	time (s)
Fig. 5 (1)	481×321	1.453
Fig. 5 (2)	481×321	1.829
Fig. 5 (3)	321×481	1.622
Fig. 5 (4)	481×321	1.560

Table 2. Computational efficiency of the examples in Fig. 5.

4. CONCLUSIONS

In this paper, we propose an iterative refinement framework with geodesic distance enhancement active contour model. The iterative seeds refinement strategy gradually update the reliable segment part and use them as guidance to segment the unreliable part. With these two novel strategies, the proposed method can handle images with complex textures and limit quantity of input seeds. Experiments on public image segmentation dataset and natural images indicate that our proposed method can achieve high quality segmentation.

5. ACKNOWLEDGMENT

6. REFERENCES

[1] Y.Y Boykov and M-P Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in nd images," in *ICCV*. IEEE, 2001, vol. 1, pp. 105–112.

[2] C Rother, V Kolmogorov, and A Blake, "Grabcut: Interactive foreground extraction using iterated graph cuts," in *TOG*. ACM, 2004, vol. 23, pp. 309–314.

[3] Leo Grady, "Random walks for image segmentation," *TPAMI*, vol. 28, no. 11, pp. 1768–1783, 2006.

[4] J Zhang, J Zheng, and J Cai, "A diffusion approach to seeded image segmentation," in *CVPR*. IEEE, 2010, pp. 2125–2132.

[5] X Bai and G Sapiro, "A geodesic framework for fast interactive image and video segmentation and matting," in *ICCV*. IEEE, 2007, pp. 1–8.

[6] A Criminisi, T Sharp, and A Blake, "Geos: Geodesic image segmentation," in *ECCV*, pp. 99–112. Springer, 2008.

[7] A.K Sinop and L Grady, "A seeded image segmentation framework unifying graph cuts and random walker which yields a new algorithm," in *ICCV*. IEEE, 2007, pp. 1–8.

[8] X Bresson, S Esedolu, P Vandergheynst, J-P Thiran, and S Osher, "Fast global minimization of the active contour/snake model," *Journal of Mathematical Imaging and vision*, vol. 28, no. 2, pp. 151–167, 2007.

[9] Tony F. Chan and Luminita A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[10] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Comm. Pure Appl. Math.*, vol. 42, no. 5, pp. 577–685, July 1989.

[11] T.N.A Nguyen, J Cai, J Zhang, and J Zheng, "Robust interactive image segmentation using convex active contours," *TIP*, vol. 21, no. 8, pp. 3734–3743, 2012.

[12] Michael Kass, Andrew P. Witkin, and Demetri Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[13] Vicent Caselles, Ron Kimmel, and Guillermo Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, 1997.

[14] A Levin, D Lischinski, and Y Weiss, "Colorization using optimization," *TOG*, vol. 23, no. 3, pp. 689–694, 2004.

[15] D Lischinski, Z Farbman, M Uyttendaele, and R Szeliski, "Interactive local adjustment of tonal values," in *TOG*. ACM, 2006, vol. 25, pp. 646–653.

[16] L Xu, Q Yan, and J Jia, "A sparse control model for image and video editing," *TOG*, vol. 32, no. 6, pp. 197, 2013.

[17] Tom Goldstein, Xavier Bresson, and Stanley Osher, "Geometric applications of the split bregman method: Segmentation and surface reconstruction," *J. Sci. Comput.*, vol. 45, no. 1-3, pp. 272–293, 2010.

[18] Tom Goldstein and Stanley Osher, "The split bregman method for l1-regularized problems," *SIAM J. Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[19] "Demo," youtu.be/NUGXALNgIH0.

[20] "Supplimental materials," home.ustc.edu.cn/∼yao2peng/icip2014.

[21] Olivier Duchenne, J-Y Audibert, Renaud Keriven, Jean Ponce, and Florent Ségonne, "Segmentation by transduction," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8.