

Auto-marking Image Segmentation Based Manifold Ranking

X.H. Zeng

Chongqing Key Lab of Computational Intelligence
Chongqing, China
College of Computer Science and Technology
Chongqing University of Posts and Telecommunications
Chongqing, China

S.W. Zhu

Chongqing Key Lab of Computational Intelligence
Chongqing, China
College of Computer Science and Technology
Chongqing University of Posts and Telecommunications
Chongqing, China

R.H. Yi

Chongqing Key Lab of Computational Intelligence
Chongqing, China
College of Computer Science and Technology
Chongqing University of Posts and Telecommunications
Chongqing, China

S.S. He

Chongqing Key Lab of Computational Intelligence
Chongqing, China
College of Computer Science and Technology
Chongqing University of Posts and Telecommunications
Chongqing, China

Abstract—Interactive image segmentation requires adjusting label information manually which will lead to tedious marking process. We propose an auto-marking image segmentation method. It can obtain the object prior and background prior automatically. We segment an image into superpixels (regions) and get the object saliency map via the manifold ranking in the guide of background prior information, then we choose part of the superpixels with higher saliency values as object marked seeds, and select the background marked seeds with the combination of background prior and the result of manifold ranking, thus obtaining the final image segmentation through maximal similarity based region merging. Experimental results on images with single object and similar adjacent objects show that the proposed algorithm can automatically add the correct label information, and can obtain segmentation accuracy that is better than saliency-seeded region merging (SSRMf) algorithm, while is more convenient than interactive segmentation by avoiding manual operations.

Keywords—image segmentation; manifold ranking; saliency; background prior; auto-marking

I. INTRODUCTION

Image segmentation [1] is a basic computer vision technology. The information in images including complexity and diversity makes image segmentation become a hot and difficult problem. In recent years, Interactive image segmentation based on part prior information of background or object is more reliable, such as Li Yin et al proposed Lazy Snapping [2] using K-means clustering algorithm on the foreground and background cluster analysis. Especially, Jifeng Ning et al proposed interactive image segmentation algorithm by maximal similarity based region merging (MSRM) [3], which can adaptively merge pre-segmentation regions with the highest similarity in the guide of label information.

Although the interactive segmentation result is accurate,

when processing massive images, manual marking will render segmentation inefficient. Furthermore, we often fail to get the ideal label for once, and have to adjust label laboriously. Inspired by saliency detection [4, 5] which can capture object region, we can use it to obtain object prior information. The paper [12] is based on the idea, but the auto-marking information is biased. Recently, background prior is used for saliency detection, such as C. Yang's saliency detection algorithm [6]. It sorts the pre-segmentation regions using manifold ranking [7] by considering the combination of background and foreground clues. We can benefit from it.

In order to make interactive segmentation more convenient, we propose an auto-marking image segmentation algorithm based manifold ranking (AISMR). Some improvements are made in this paper. Firstly, we obtain object seeds automatically rather than manually by saliency detection. Secondly, we mark the background seeds correctly by the principle of the combination of background priors and the result of manifold ranking, through which we can effectively improve the accuracy of image segmentation.

II. AUTO-MARKING IMAGE SEGMENTATION BASED MANIFOLD RANKING

Some pre-segmentation methods which divide the image into many small regions are used in image processing. We use the simple linear iterative clustering (SLIC) algorithm [8] for image pre-segmentation. The image is segmented into many small regions called superpixels, which makes it easier in later ranking and merging process.

Saliency region of the image can be used as the object prior information. The paper [6] has shown that manifold ranking is an efficient method for image saliency detection. It calculates the ranking scores using global structure of manifold among regions, which can reveal correlation among regions appropriately, and the higher scores will be considered

as salient object regions.

Therefore, we use manifold ranking for superpixels to get salient object regions, and view the part of image which have higher saliency value as object marked seeds, then use the background prior (image boundary superpixels as background marked seeds) to remove the superpixels with higher saliency value from those sorted by manifold ranking, so the rest superpixels will be viewed as background marked seeds among boundary superpixels. Finally, we merge superpixels by maximal similarity region merging criterion [3] to obtain object segmentation. Above is the auto-marking image segmentation based manifold ranking (AISMR) we proposed. The segmentation model is shown in Fig.1:

A. Obtain Label Information by Manifold Ranking

We obtain superpixels (regions) by SLIC, and then construct a weighted graph, in which each node represents a superpixel. When initializing, the default background superpixels are endowed with positive sorting scores, and the scores of remaining candidate superpixels are 0. Every node in the network will spread its score to its neighbor nodes before the network reaches equilibrium. All the scores of candidate superpixels are used to sort the superpixels, through which we get an object saliency map. The algorithm is described as follows:

Given a superpixels set $X=\{x1, \dots, xk, xk+1, \dots, xn\} \in R^{m \times n}$ of an image, where $\{x1, \dots, xk\}$ are marked boundary superpixels, $\{xk+1, \dots, xn\}$ are unmarked superpixels which need to be sorted. Let $f: X \rightarrow R^{m \times n}$ denote a ranking function which assigns a ranking value fk to each superpixel xk , and f can be seen as a vector $f=[f1, \dots, fn]^T$. Let $y=[y1, y2, \dots, yn]^T$ denote an indication vector. If xk is a marked superpixel, then $yk=1$, otherwise $yk=0$.

We can view each superpixel as a node and construct a nearest neighbor graph $G(V,E)$, where V denote all the nodes, E denote the weights between nodes. These weights are computed by the distance in color space. Each node is not only connected to its neighbor nodes, but also to the nodes sharing same boundaries with its neighbor node. The weight is defined by:

$$w_{mn} = \begin{cases} \exp(-\frac{\|C_m - C_n\|}{\sigma^2}), & n \in (A_m \cup S_{A_m}) \\ 0, & n \notin (A_m \cup S_{A_m}) \end{cases} \quad m, n \in V \quad (1)$$

where the C_m and C_n denote the mean of the superpixels corresponding to two nodes m, n in the CIE LAB color space [11], and σ is a constant that controls the strength of the weight. Where A_m denotes adjacent superpixels of region m , S_{A_m} denotes A_m 's adjacent superpixels. Then we consider four boundary nodes, any pair of boundary nodes is considered to be adjacent. Through this way we can get more accurate measurement of the similarity between superpixels, which is effective for saliency detection [6].

We get the superpixels weight matrix W using Eq.1, then calculate the diagonal matrix D whose elements are the sum of each row elements in W . We use the unnormalized Laplace

matrix to calculate ranking vector f . It can get better result than normalized Laplace matrix [6]. The f can be written as:

$$f = (D - \alpha W)^{-1} y \quad (2)$$

where α controls the scores both come from the prior and the structure of the adjacent nodes, which contribute to the final sorting scores. The larger value of α means the greater proportion contribution coming from adjacent nodes scores ($\alpha \in [0, 1)$). The inverse matrix $(D - \alpha W)^{-1}$ is an affinity matrix which can capture the spatial relation among superpixels. It means that the similarity between superpixels increase when their spatial distance decrease.

The four boundaries in an image are usually the background regions. We can use this background prior to acquire the saliency scores of object regions. $\{X1, X2, X3, X4\}$ denotes the set of superpixels on each side of image, and we regard each element as background marked queries respectively. First of all, we compute the sorting scores $f1$ of all the superpixels using Eq.2. When $X1$ is the background marked queries, we get the sorting scores vector f_1^* according to normalized $f1$, then acquire the object's saliency scores $1 - f_1^*(x_k)$ $x_k \in X$. Next, we regard $\{X2, X3, X4\}$ as background marked queries respectively which can acquire $1 - f_2^*(x_k)$, $1 - f_3^*(x_k)$, $1 - f_4^*(x_k)$ corresponding to object's saliency scores. To weaken the interference of sorting scores where contains the object regions in boundary, we render four object's saliency scores using mutual dot product (\odot), then normalize it to obtain the ultimate object's saliency scores $F(x_k)$.

$$F(x_k) = (1 - f_1^*(x_k)) \odot (1 - f_2^*(x_k)) \odot (1 - f_3^*(x_k)) \odot (1 - f_4^*(x_k)), \quad x_k \in X \quad (3)$$

In order to better observe the saliency object and select threshold of seeds, the sorting scores $F(x_k) \times 255$ are mapped to a gray image which can get the saliency map $S(x_k)$ $x_k \in X$. Although the saliency map can't present the overall outline of object completely, it includes important saliency features of object regions. We can easily get the correct marked seeds by a simple operation. The correct marked seeds are very important for the following merging process.

B. Marking Object Seeds and Background Seeds Automatically

Combined with the background prior and saliency map using manifold ranking, our method can obtain the object and background marked seeds where the image with a single object or similar adjacent objects more easily.

Based on the fact that object has higher saliency value, we can obtain the object marked seeds by automatically setting the threshold O :

$$O = \{(m, n) | P(S(x_k) > S_g) = \beta\}, \quad x_k \in X \quad (4)$$

where (m, n) denotes pixel coordinates, S_g denotes a

threshold which is 200 to get more precise mark. $P(S(x_k) > S_g)$ denotes superpixels' saliency values which are larger than S_g , and we regard it as an object label β in saliency map, where β is an object label value.

We use the background prior again to obtain background marked seeds. Taking the superpixels of boundary regions as background label B . B can be written as:

$$B = \{(m, n) | P(X_1 \cup X_2 \cup X_3 \cup X_4) = \gamma\} \quad (5)$$

where $\{X_1, X_2, X_3, X_4\}$ denote set of four boundary superpixels regions. $P(Q)$ denotes all the pixels in a set of superpixels Q , where $Q = X_1 \cup X_2 \cup X_3 \cup X_4$, γ is a background label value.

The boundary regions may contain part of the object. If all of the boundary superpixels are background marked seeds, the label may cause the error of segmentation results, so we must remove the part of the superpixels X_o belonging to the object in the boundary region in Eq.5. We obtain the boundary superpixels' saliency value sorting sequence $S(S(x_1), S(x_2), \dots, S(x_l))$ by previous manifold ranking. As part of the object has high saliency value, we can easily remove X_o among boundary superpixels.

$$X_o = \{x_k | S(x_k)_{k=1,2,\dots,l} > S_o\}, x_k \in B \quad (6)$$

where s_o denotes a threshold which is 1 to get more precise mark. Then, we choose the background label B' again, which can be written as:

$$B' = \{(m, n) | P((X_1 \cup X_2 \cup X_3 \cup X_4) - X_o) = \gamma\} \quad (7)$$

We can obtain the correct background marked seeds using Eq.7. It means the background regions without the object regions in boundary.

C. Superpixels merging

We will add object and background marked seeds which have got above to the preprocessing image automatically, and then obtain the final object segmentation image by maximal similarity based region merging.

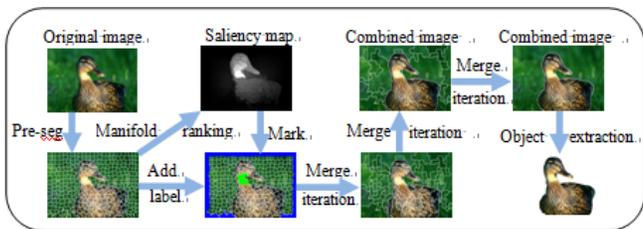


FIGURE I. AISMR MODEL

D. Auto-Marking Segmentation Algorithm

Input: Original color image

Output: Segmentation object

Step1: Image pre-segmentation. Use SLIC algorithm to obtain superpixels set $X = \{x_1, \dots, x_k, x_{k+1}, \dots, x_n\}$.

Step2: Superpixels ranking. Construct a nearest neighbor graph $G(V, E)$ and acquire its weight matrix according to Eq.1. Use background priors to get superpixels sorting score vector $f = [f_1, f_2, \dots, f_n]^T$ according to Eq.2.

Step3: Obtaining the saliency map. Get four object's saliency scores $1 - f_i^*(x_k)$ $i=1, 2, 3, 4$ according to background prior $\{X_1, X_2, X_3, X_4\}$ in sequence. Obtain the object's scores $F(x_k)$ according to Eq.3. Acquire the saliency map $S(x_k)$ by $F(x_k) \times 255$.

Step4: Marking the object seeds. View the superpixels which satisfy the conditions $S(x_k) > S_g$ as object marked seeds O in the saliency map $S(x_k)$.

Step5: Marking the background seeds. Regard $\{X_1, X_2, X_3, X_4\}$ as background marked seeds B . Get the sorting sequence of $S(S(x_1), S(x_2), \dots, S(x_l))$ using the previous saliency map $S(x_k)$. Obtain X_o through $S(x_k)_{k=1,2,\dots,l} > S_o$. Reset the background marked seeds B' .

Step6: Merging adjacent superpixels. Calculate similarity $\rho[9]$ between the superpixel and its adjacent superpixels by Bhattacharyya distance[10]. Merge the adjacent superpixels with the maximal similarity region merging criterion using label information.

III. EXPERIMENTAL RESULTS AND ANALYSIS

To prove the effectiveness of our algorithm, we select some images such as plants, animals and humans on the Internet to have a test in Matlab2012b. Fig.2 shows the comparison results of manual reference segmentation, SSRMf [12], MSRM [3], and our algorithm, respectively.



(a) Image (b) Reference (c) SSRMf (d) MSRM (e) Ours

FIGURE II. IMAGE SEGMENTATION RESULTS OF DIFFERENT METHODS

The experimental results are in Fig.2. The Fig.2 (b) is used as the reference of segmentation object. In Fig.2(c), SSRMf can segment the object including starfish, but the segmentation object contains many background regions because the marked seeds are not accurate. In the image of deer, SSRMf can't segment the object completely. The main reason is that the marked seeds do not cover the key features of the object or background marked seeds which contain a small portion of the object region. In Fig.2 (d), MSRM can obtain the object segmentation completely, but it needs manual marking to get better results. In Fig.2 (e), our algorithm can obtain better object segmentation without manual marking. When observing

the image segmentation directly, the accuracy of our algorithm is higher than SSRMF and the effect of it is similar with MSRM.

The TPR and FPR value [3] are used as segmentation evaluation index, where the TPR value represents the correct segmentation ratio; the FPR value represents the error of segmentation ratio. Fig.3 illustrates the TPR value and FPR value of three algorithms above.



FIGURE III. (A) THE TPR VALUES (B) THE FPR VALUES

An image segmentation object with higher TPR value and lower FPR value is viewed as a better one. In Fig.3, SSRMF's TPR and FPR value are very high while MSRM's TPR value is the highest and the FPR value is the lowest in most images. AISMR's TPR value is slightly lower than MSRM's and AISMR's FPR value is slightly higher than MSRM's. Compared with these two algorithms, our algorithm is better than SSRMF and can obtain the similar results with MSRM. Experimental results show that our method can ensure the accuracy and enhance the efficiency at the same time.

IV. CONCLUSION

Correct prior information is the key part of an interactive image segmentation algorithm. When the label information can cover the key features of the object or background region, the method adding the label information automatically will make the operation more convenient and achieve the similar effect of MSRM. Enlightened by this idea we proposed an efficient algorithm. It has achieved effective segmentation results on images of single object or similar adjacent objects. However, if color features are similar between object and the background region in an image, it will be difficult to extract the object from the background region. In the future, we will focus on improved manifold ranking algorithm that can obtain more accurate label information and achieve reliable segmentation results.

ACKNOWLEDGEMENT

This work was partly supported by the National Natural Science Foundation of China (No. 61075019, 61379114).

REFERENCES

[1] Lucchesez L, Mitray S K. Color image segmentation a state of the art survey [J]. *Image Processing Vision and Pattern Recognition*, 67(2): 207-221, 2001.

[2] Y. Li, J. Sun, C.K. Tang, et al. Lazy snapping. *ACM Transaction on Graphics*, 23(3):203-308, 2004.

[3] J.F. Ning, L. Zhang, D. Zhang and C.K. Wu. Interactive image segmentation by maximal similarity based region merging. *Pattern Recognition*, 43(2): 445-456, 2010.

[4] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE TPAMI*, 20(11):1254-1259, 1998.

[5] J. Harel, C. Koch, P. Perona. Graph-Based Visual Saliency. *Advances in Neural Information Processing Systems*, vol. 19, pp. 545-552, 2007.

[6] C. Yang, L. Zhang, H. Lu, X. Ruan and M.H. Yang. Saliency detection via graph-based manifold ranking. *In CVPR*, 2013.

[7] D. Zhou, J. Weston, A. Gretton, O. Bousquet, and B. Scholkopf. Ranking on data manifolds. *In NIPS*, 2004.

[8] R. Achanta, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. Slic superpixels. *Technical report. EPFL*, Tech.Rep. 149300, 2010.

[9] K. Fukunaga. Introduction to Statistical Pattern Recognition. *Academic Press*, 1990.

[10] T. Kailath. The divergence and Bhattacharyya distance measures in signal selection. *IEEE Transactions on Communications Technology* 15 (1): 52-60, 1967.

[11] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-tuned salient region detection. *In CVPR*, 2009.

[12] J.X. Li, R.N. Ma, J.D. Ding. Saliency-Seeded Region Merging: Automatic Object Segmentation. *In ACPR*, 2011