

GRAPH BASED SEGMENTATION WITH MINIMAL USER INTERACTION

Huaizhong Zhang, Ehab Essa, and Xianghua Xie*

Computer Science Department, Swansea University,
Swansea SA2 8PP, United Kingdom

*<http://csvision.swan.ac.uk/>

ABSTRACT

In this paper, we present a graph based segmentation method that only requires a single point from user initialization. We incorporate a new image feature into the segmentation scheme. It is derived from a vector field that takes into account gradient vector interactions across the image domain, and has the simplicity of edge based features but also proves to be a useful region indication in two-level segmentation. Effective vector field diffusion is proposed to deal with excessive image noise. Based on a single user point we unravel the image and transfer the object segmentation into a height field segmentation in polar coordinates, which in effect imposes a star shape prior. The search of a minimum closed set on a node weighted, directed graph produces the segmentation result. Comparative analysis on real world images demonstrates promising performances of the proposed method in segmentation accuracy and its simplicity in user interaction.

Index Terms— Graph cut, interactive segmentation, optimal surface segmentation, star graph

1. INTRODUCTION

Interactive image segmentation is of increasing interest in recent years due to their practical use in various applications, e.g. [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11]. Among these methods, graph cut technique [4, 6] is a common approach to incorporating user interaction in two-level segmentation due to its guaranteed polynomial time efficiency in reaching global minimum. The min-cut/max-flow algorithms, widely used in realizing the cut [4, 6, 12, 13, 14], are formulated by encoding the user interaction term (data) and the regularization term (prior) where the augmenting path algorithm is found superior for computer vision problems [13]. User interactions can help model and update the object/background information so that a stable representation of the objects of interest can be obtained and accurate segmentation is therefore achieved. However, conventional graph cut methods are with little shape constraints and greatly depend on user interactions; thus, the extracted shapes are sensitive to initial seed selections. In addition, extensive manual interferences may significantly reduce the efficiency of segmentation. Considerable effort has been made to deal with these limitations,

particularly the use of shape prior for reducing user interactions. Object specific shape prior [15] has been shown to be a powerful approach whereas incorporating generic priors for general image segmentation remains a challenging task. An elliptical shape prior in [16] is used to iteratively refine object extraction while a blob-like shape prior is adopted in [17, 7]. However, these shape priors are too restrictive for general image segmentation. Vicente *et al.* in [8] extended the grab cut [3] to improve its performance towards thin and elongated structures, at the cost of much greater user interaction. Embedding user input into a distance function to regularize graph cut [18] has also been found to improve the accuracy, the degree of which may largely depend on its user initialization. Recently, the star-like shape assumption is proposed in [10] to provide a generic shape prior for graph cut segmentation. Only a single user input is imposed and a global optimizer is obtained subject to the star-like prior. To avoid bias towards small segments, a length-based “balloon” force term is used in the cost function, which can be intractable. Very recently, this method is extended to deal with more complex shape by allowing more user input [11].

In this work, we adopt a similar assumption as in star graph [10], that is given a star point, often the center of an object, the object boundary does not occlude itself from the star point. This assumption does not apply to all objects, but is applicable to a large number of real world objects as demonstrated in [10]. Note, the objects do not have to be convex, and “star” is a more generic term used to refer to this shape prior than assuming the objects have to be star-like. But, instead of constructing the graph in the original image domain, we unravel the image into polar coordinates and carry out a different graph construction. The star prior is hence translated to the assumption that object boundary in the polar coordinates is unfolded, but by doing so, we remove the bias towards shorter cut and avoid the dedicated process to work out the graph path from star point to each image pixel. This transformation also allows us to apply the optimal surface graph construction [19], whose global optimality is guaranteed and it can be searched without user interference. Moreover, we incorporate a novel image feature to the cost function, instead of merely image intensity or local gradient. The image feature is derived from the gradient vector interaction across the

image domain and possesses the characteristics of regional features. Another benefit is that diffusion can be effectively taken place in its vector field to deal with image noise. We show this image feature, by combining with optimal graph cut and star prior, provides promising segmentation performances with minimal user interaction, that is a single user point. In the following section, we describe the details of the proposed approach both in modeling and implementation. The experimental results are presented in Section 3 and the paper is concluded in Section 4.

2. PROPOSED METHOD

2.1. Image feature

In graph cut, the cost function can be generally categorized as edge based and region based. Edge based cost functions assume that object boundary is largely collocated with image intensity discontinuity, and typically use derivatives of image intensity function as a local estimation of likelihood of an object boundary. Region based ones are usually non-edge based, e.g. piecewise constant assumption. Quite often, image intensity values are directly used in general image segmentation. Although graph cut algorithms provide global optimality in two-level segmentation, a reliable but also generic image feature that does not assume strong image prior is desirable for general segmentation that is useful for, for instance, object recognition. We consider intensity discontinuity perhaps is the least constrained and most widely applicable object boundary estimation. Its performance can be easily compromised by image noise, smooth varying intensity at object boundary, and so on. These shortcomings are essentially because it is a local measurement and it does not take into account interactions among image gradient vectors. As an example, a region with relatively large image gradient magnitude by varying gradient directions suggests that it is unlikely a location of object boundary, despite their large magnitude. On the contrary, weak gradient vectors that are aligned to each other suggest greater likelihood of object boundary than what the magnitude itself suggests. Hence, we present a gradient vector field that is a result of global interactions among original image gradient vectors, and we show that its circulation density can be used as a reliable image feature for graph cut. The zero-crossings of this circulation density provides a better indication of the location of object boundary, and the magnitude of oscillation at zero-crossings indicates the strength of object boundary presence. The signs (positive and negative) of circulation density actually indicate the foreground and background. The derived gradient vector can also be diffused to produce more coherent circulation density. The image feature is directly derived from edge based assumption, but resembles closely to region based methods.

Let $\nabla_i I = f\hat{I}_x$ and $\nabla_j I = f\hat{I}_y$ denote the two components of the image gradient ∇I in the image coordinates (i, j) , respectively, i.e. $\nabla I = (\nabla_i I, \nabla_j I)^T$ where f is edge

map (magnitude). We carry out the convolution computation on both components with the kernel $k(\mathbf{x}) = m(\mathbf{x})$. Moreover, we choose the magnitude function m as an inverse of distance, i.e. $m(r) = 1/r^\zeta$ with $\zeta = 1$. Since we further compute the spatial derivatives of the convolution results, the spatial decay is actually raised to power of two, i.e. $\zeta = 2$. Thus, the result of this convolution process can be expressed as:

$$\begin{cases} E_i(\mathbf{x}) = \nabla_i I * k(\mathbf{x}) = \sum_{s \neq \mathbf{x}} \frac{\nabla_i I(s)}{R_{\mathbf{x}s}} = \sum_{s \neq \mathbf{x}} f(s) \frac{\hat{I}_x(s)}{R_{\mathbf{x}s}}, \\ E_j(\mathbf{x}) = \nabla_j I * k(\mathbf{x}) = \sum_{s \neq \mathbf{x}} \frac{\nabla_j I(s)}{R_{\mathbf{x}s}} = \sum_{s \neq \mathbf{x}} f(s) \frac{\hat{I}_y(s)}{R_{\mathbf{x}s}}, \end{cases} \quad (1)$$

where $R_{\mathbf{x}s}$ is the distance from x to s in the image plane and $\mathbf{E} = (E_i, E_j)$ denotes the resulting gradient convolution field. Due to the smoothing effect when applying the kernel function, the original image gradient vectors have extended their influence from immediate vicinity of edge pixels to much larger neighborhood. In fact, the computation in (1) is across the whole image domain.

Next, we compute the circulation density, i.e.

$$B = \nabla \cdot \mathbf{E}(\mathbf{x}) = \nabla \cdot (E_i, E_j) = \nabla \times (-E_j, E_i). \quad (2)$$

It can be shown that this circulation density has an intrinsic link to the magnetic field used in the MAC model [2] in a variational framework. Specifically, when $\zeta = 1$, B is equivalent to the third and only effectively component of the magnetic field in the MAC model. Hence, the positive and negative values of this circulation density indicate foreground/background and background/foreground. The zero crossings of the circulation density would indicate the location of object boundaries. In particular, the proposed circulation density method is a generalization of the effective component used in MAC. Moreover, we can refine the computation of this circulation density by performing efficient Laplacian diffusion in the extended gradient vector field to overcome, for instance, noise interference.

Note that the gradient vector field is actually along the edge direction so substantial diffusion in the components can result in significantly improved boundary description. There are various diffusion strategies for this smoothing task. For implementation convenience and less parameter intervention, we use an isotropic/Laplacian diffusion scheme in this paper, which is carried out by solving the following Euler equations.

$$\begin{cases} \frac{\partial}{\partial y} \mathcal{E}_i = p(E_i) \Delta \mathcal{E}_i - q(E_i) (\mathcal{E}_i - E_i), \\ \frac{\partial}{\partial i} \mathcal{E}_j = p(E_j) \Delta \mathcal{E}_j - q(E_j) (\mathcal{E}_j - E_j), \end{cases} \quad (3)$$

where $\mathcal{E}_i(0, \mathbf{x}) = E_i(\mathbf{x})$, $\mathcal{E}_j(0, \mathbf{x}) = E_j(\mathbf{x})$, and $p(y)$ and $q(y)$ are given as:

$$p(y) = \exp(-|y|f/K), q(y) = 1 - p(y), \quad (4)$$

where $f = |\nabla I|$ and K is a constant.

2.2. Graph construction

The image is unraveled from the original Cartesian coordinates to polar coordinates based on a user specified origin,



Fig. 1: Top row: original image, MAC [2] result, and $s - t$ cut [4] result. Bottom row: star graph [10] result, result of the proposed method without image feature diffusion, and result of the proposed method with diffusion.

which is usually the center of the object of interest. We thus assume that the object boundary intersects with each column of pixels once and once only, which is equivalent to the star graph assumption [10] that object boundary does not occlude itself from the single user point in the Cartesian coordinates. However, transforming the image into polar coordinates eliminates the need for a length penalty term in the cost function to avoid bias towards shorter cut. It also provides a regular graph grid that is easy to work on, in contrast to the nonuniform graph structure in star graph method [10]. It is hence natural to adopt the optimal surface method, proposed in [19], to construct the graph as it guarantees global optimal.

For the constructed graph $G = \langle V, E \rangle$, each node $V(x, y)$ corresponds to a pixel in 2D image $I(x, y)$. There are two types of arcs included in the graph G : intra-column arcs and inter-column arcs. For intra-column, every node $V(x, y)$ where $y > 0$ has a directed arc to the node $V(x, y - 1)$. For inter-column, each node $V(x, y)$ links to a node $V(x + 1, \max(0, y - \Delta))$ with a directed arc, where Δ is a smoothing factor. Similarly, node $V(x + 1, y)$ is established to connect $V(x, \max(0, y - \Delta))$. The last row of the graph is to connect each other of the nodes so as to keep a closed graph. Fig. 2(a) provides an illustration of this directed graph on two columns. Fig. 2(b) shows the unravelled image given in Fig. 1. Image cost is then assigned to the nodes of this directed graph, which is then transformed to an edge weighted graph with the upper interface of its minimum closed set from bottom of the graph corresponds to an optimal cut.

2.3. Cost function and graph cut

Boundary based cost function is used in our solution scheme, whose energy is described as $E = \sum_{V \in S} \hat{C}(x, y)$ where $\hat{C}(x, y)$ denotes the normalized cost function ($\hat{C}(x, y) \in [0, 1]$) and S is a path in the derived directed graph. As aforementioned, the zero-crossings of the circulation density feature computed from gradient vector interaction indicate the location of object boundary. The degree of circulation

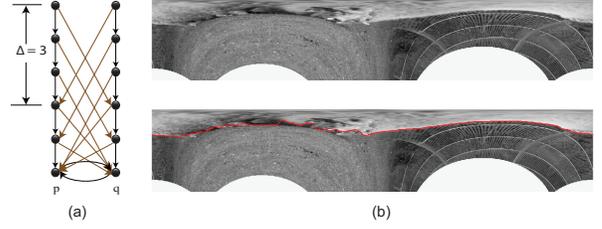


Fig. 2: (a) An illustration of inter-column and intra-column arcs. (b) An example unravelled image and its segmentation result.

density oscillation at the zero-crossing suggests the strength of object boundary. A direct assignment of circulation density to the graph as nodal cost would be inappropriate, since the minimum or maximum of circulation density is not an indication of either location or strength of object boundary. However, a simple transformation, for instance computing its gradient magnitude, can be applied. Since the decay of circulation density from object boundary is exponential, a log transformation can be added in order to avoid extreme values to be assigned to the graph, i.e. $C = -\log|\nabla B|$. Each graph node is then weighted by a value representing its rank to be selected in the minimum closed set graph where the arc costs between graph nodes are infinitive. The weight of each graph node is assigned as follows:

$$\omega(x, y) = \begin{cases} C(x, y) & \text{if } y = 0, \\ C(x, y) - C(x, y - 1) & \text{otherwise.} \end{cases} \quad (5)$$

where C is the cost function and ω the weight for each node in the directed graph, which the weighted nodes can be further decomposed into nonnegative and negative sets. Accordingly, the minimum closed set can be obtained by the $s - t$ cut method where the source s is to connect each negative node and every nonnegative node is connected to the sink t . The interface, corresponding to the object boundary, is presented by solving this $s - t$ cut problem. Fig. 2(b) shows an example of segmentation in the polar coordinates and its result in Cartesian coordinates is given in Fig. 1.

3. EXPERIMENTAL RESULTS

The proposed method was tested on various real world images and was compared against a number of segmentation techniques, including MAC model [2], $s - t$ graph cut [4], and star graph [10]. Due to lack of space, the comparison with MAC and conventional graph cut is briefly presented, and our method is mainly compared against star graph method, which is mostly close to the proposed method. In all cases, the star graph and the proposed method were using the same user initializations. We found that the balloon force in the star graph method has a significant impact on its segmentation result. Hence, multiple runs were carried out for star graph, follow-

ing the careful parameter selection according to the suggestion by the authors in [10], and the best results were presented here.

In Fig. 1, the MAC model with the diffused magnetic field produced a reasonable result, but it under-segmented the object due to large variations at the bottom half of the object. The fact that a gradient descent optimization was used also suggests that it may be trapped at a local minima. The $s - t$ cut required the user to specify foreground and background in order to assign cost function and establish terminal links. Its result however was not satisfactory due to varying image intensity. Star graph method, actually affected by the textured background, over-segmented the object. On the other hand, the proposed method without diffusion achieved reasonable result, which was further improved by performing image feature diffusion. The user initialization for both star graph and the proposed method is minimal, i.e. a single point.

Fig. 3 provides typical comparative results of the proposed method against the star graph method. The original images are shown in the first column (a), followed by results of star graph (b) and results of the proposed method (c). The testing images have varying degrees of difficulties, such as weak edges, nonuniform intensity, textured appearance, and cluttered background. Albeit careful parameter tuning with the star graph method, it tends to under-segment objects in the presence of weak edges and over-segment when objects are textured. In fact, this is due to the use of the intractable “balloon” force that is for dealing with the bias issue in the star graph method. In contrast, by effectively taking the image feature, the proposed method achieved consistently superior performance with the same user initializations. For further demonstrating the effectiveness of the proposed method, another eight examples using the proposed method are presented in Fig. 4. Note that our method had some difficulties in the last two examples in Fig. 4 because their thin and long structures posed challenge to our image unraveling strategy.

4. CONCLUSIONS

We propose a graph based segmentation that requires minimal user interaction. Its cost function is based on a novel image feature that is derived from global interactions of image gradient vectors. Diffusion scheme could be applied to further refine the features in order to produce more coherent segmentation. Preliminary comparative analysis on real world images showed promising performances. The method may be further extended to deal with long, thin structures, e.g. the last two examples shown in Fig. 4. Combining Dijkstra graph to extend the segmentation method may be able to tackle such issues. In addition, considering the effectiveness of segmenting the images with textured background, the proposed method may be further applied to deal with some segmentation difficulties in specific medical image modalities such as Optical Coherent Tomography, which the traditional methods are often stuck in serious artefacts/noise in the images.

(a) Original Image (b) Star Shape Method (c) Proposed Method

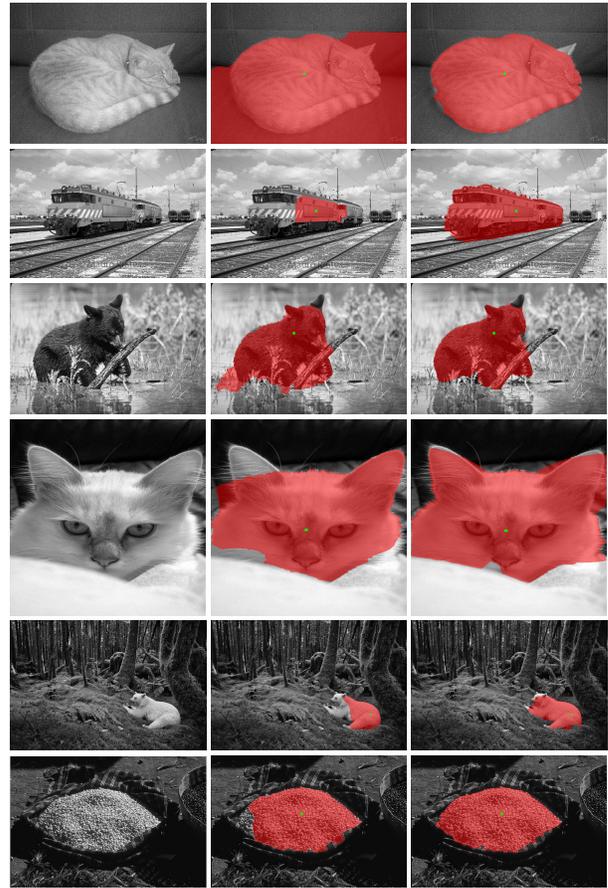


Fig. 3: Comparison results for various images. Column (a): Original Image; Column (b): Results using star graph [10]; Column (c): Results using the proposed method. The green dots indicate the user initializations.

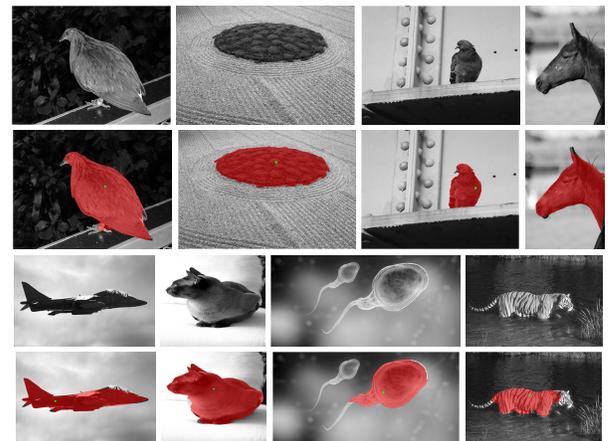


Fig. 4: Further examples of the proposed method.

5. REFERENCES

- [1] E. Mortensen and W. Barrett, "Interactive segmentation with intelligent scissors," *GMIP*, vol. 60, pp. 349–384, May 1998.
- [2] X. Xie and M. Mirmehdi, "MAC: Magnetostatic active contour," *T-PAMI*, vol. 30, no. 4, pp. 632–646, 2008.
- [3] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut: interactive foreground extraction using iterated graph cuts," in *SIGGRAPH*, 2004, pp. 309–314.
- [4] Y. Boykov and M. Jolly, "Interactive graph cuts for optimal boundary and region segmentation," in *ICCV*, 2001, pp. 105–112.
- [5] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *PAMI*, vol. 23, pp. 1222–1239, Nov. 2001.
- [6] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient n-d image segmentation," *IJCV*, vol. 70, pp. 109–131, Feb. 2006.
- [7] P. Das, O. Veksler, and Y. Boykov, "Semiautomatic segmentation with compact shape prior," *CRV*, vol. 62, pp. 28–36, January 2006.
- [8] S. Vicente, V. Kolmogorov, and C. Rother, "Graph cut based image segmentation with connectivity priors," in *CVPR*, 2008.
- [9] L. Gorelick, A. Delong, O. Veksler, and Y. Boykov, "Recursive mdl via graph cuts: Application to segmentation," in *ICCV*, 2011.
- [10] O. Veksler, "Star shape prior for graph-cut image segmentation," in *ECCV*, 2008, pp. 454–467.
- [11] V. Gulshan, C. Rother, A. Criminisi, A. Blake, and A. Zisserman, "Geodesic star convexity for interactive image segmentation," in *CVPR*, 2010, pp. 3129–3136.
- [12] A. Goldberg and R. Tarjan, "A new approach to the maximum-flow problem," *Journal of the Association for Computing Machinery*, vol. 35, pp. 921–940, 1988.
- [13] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *PAMI*, vol. 26, pp. 1124–1137, Sept. 2004.
- [14] O. Juan and Y. Boykov, "Active graph cuts," in *CVPR*, 2006, pp. 1023–1029.
- [15] M. Kumar, P. Torr, and A. Zisserman, "Obj cut," in *CVPR*, 2005, pp. 18–25.
- [16] G. Slabaugh and G. Unal, "Graph cuts segmentation using an elliptical shape prior," in *ICIP*, 2005, pp. 1222–1225.
- [17] G. Funka-Lee, Y. Boykov, C. Florin, M. Jolly, R. Moreau-Gobard, R. Ramaraj, and D. Rinck, "Automatic heart isolation for ct coronary visualization using graph-cuts," in *ISBI*, 2006, pp. 614–617.
- [18] D. Freedman and T. Zhang, "Interactive graph cut based segmentation with shape priors," in *CVPR*, 2005, pp. 755–762.
- [19] K. Li, X. Wu, D. Chen, and M. Sonka, "Optimal surface segmentation in volumetric images—a graph-theoretic approach," *PAMI*, vol. 28, pp. 119–134, January 2006.