

A fast marine sewage detection method for remote-sensing image

Guoqiang Huan^{1,2} · Zhanjie Song¹ · Shuo Zhang^{1,3} · Jianhua Zhu⁴

Received: 7 August 2017 / Accepted: 3 January 2018 / Published online: 8 February 2018
© SBMAC - Sociedade Brasileira de Matemática Aplicada e Computacional 2018

Abstract This paper presents an effective method for marine sewage detection from a remote-sensing image. It is inspired by the Grab-Cut mechanism that iterative estimation and incomplete labeling allow a considerably reduced degree of user interaction for a given quality of result. By establishing the relationship between the color feature and the object seeds, we first model object and background with Gaussian mixture model, respectively, followed by iteratively updating the parameter of model to decline the energy function. To improve the computation efficiency, we propose to extend the region of interest as background. The proposed method accounts for not only the effect of color feature, but also the geographical information. The experimental results demonstrate that the proposed method is more reliable in marine sewage detection compared to other state-of-the-art methods.

Keywords Image detection systems · Remote sensing and sensors · Digital image processing · Gaussian mixture model · Grab-Cut mechanism

Mathematics Subject Classification 62H25 · 60J20 · 68Q87

Communicated by Cristina Turner.

✉ Shuo Zhang
shuozhang@tju.edu.cn

¹ School of Mathematics, Tianjin University, Tianjin, China

² Wuhan Space Sanjiang Laser Industry Technology Research Institute Co., Ltd., Wuhan, China

³ Department of Mathematics, Tianjin University of Finance and Economics, Tianjin, China

⁴ National Ocean Technology Center, Tianjin, China

1 Introduction

The progress of human civilization and a comfortable lifestyle are accompanied with the pollution of the air, soil, and seas. Fortunately, the use of remote-sensing satellites allows us to track the pollution timely.

Image segmentation is dividing the image into a number of procedures with consistency and non-overlapping regions. The topic of the interactive image segmentation has received considerable attention in the computer vision community in the last decades (Kolmogorov and Zabini 2004; Boykov et al. 2010; Boykov and Funka-Lea 2006). This paper is focused on how to detect the pollution region in remote-sensing (RS) images efficiently with interaction. The aim is to achieve high performance with the modest interactive effort for users. In general, the degrees of interactive effort range from editing individual pixels, at the labor-intensive extreme, to merely touching foreground and/or background in a few locations.

2 A brief review of interactive image segmentation

In the following, we categorize different methods of interactive image segmentation by their methodology and user interfaces, mainly including segmentation in discrete domain and segmentation in continuous domain.

The known appearance models typically are assumed in discrete domain of segmentation methods. The log-likelihoods of appearance are optimized in combination with some spacial regularization. This problem is relatively simple and many methods are guaranteed globally optimal results. The appearance models jointly with segmentation are estimated in continuous domain of segmentation methods. The model parameters are treated as additional variables transforming simple segmentation energies into high-order NP-hard functionals. It is known that such methods indirectly minimize the appearance overlap between the segments.

2.1 Segmentation in discrete domain

Boykov and Jolly (2001) were the first to formulate a simple generative Markov random field (MRF) model in discrete domain for the task of binary image segmentation. This basic model can be used for interactive segmentation. Given some user constraints in the form of foreground and background brushes, i.e., regional constraints, the optimal solution is computed very efficiently with graph cut. The main benefits of this approach are: global optimality, practical efficiency, numerical robustness, ability to fuse a wide range of visual cues and constraints, unrestricted topological properties of segments, and applicability to n -dimensional problems. Thanks to their work, many articles on interactive segmentation using graph cut and a brush interface were published (Boykov and Jolly 2001; Grady 2006). It also inspired the Grab-Cut system (Rother et al. 2004), which can be used to solve a more challenging problem, namely, the joint optimization of segmentation and estimation of global properties of the segments. The benefit is a simpler user interface in the form of a bounding box. Note, such joint optimization has been used in other contexts before. An example is the depth estimation in stereo images, where the optimal partitioning of the stereo images and the global properties (affine warping) of each segment are optimized jointly (Boykov and Jolly 1999). Slabaugh and Unal (2005) proposed the energy function incorporating an elliptical shape prior, which improved the accuracy of the circular object segmentation. Juan and Boykov (2006) pointed out that the key to improve the speed of interactive segmentation is to improve the efficiency of max-flow/min-cut algorithms. They designed the ActiveCuts

(AC) algorithm. Another interesting set of discrete functionals are based on ratio, e.g., area over boundary length (Kolmogorov et al. 2007).

2.2 Segmentation in continuous domain

There are very close connections between the spatially discrete MRFs and variational formulations in the continuous domain. The first continuous formulations were expressed in terms of active contours with edges (Chan and Vese 2001), related to the well-known Mumford and Shah functional (1989). The goal is to find a segmentation that minimizes a boundary (surface) under some metric, typically image-based Riemannian metric. Traditionally, techniques such as level sets were used, which, however, are only guaranteed to find a local optimum. Recently, many of these functionals were reformulated, using convex relaxation, i.e., the solution lives in the $[0, 1]$ domain, which allows to achieve global optimality and bounds in some practical cases. An example for interactive segmentation with a brush interface is (Unger et al. 2008), where the optimal solution of a weighted total variation norm is computed efficiently. Instead of using convex relaxation techniques, the continuous problem can be approximated on a discrete grid and solved optimally in global by graph cut. This can be done for a large set of useful metrics (Kolmogorov and Boykov 2005). Theoretically, the discrete approach is inferior, since the connectivity of the graph has to be large to avoid metrication artifacts. In practice, however, artifacts are rarely visible when using a geodesic distance.

2.3 Proposed method: improved Grab-Cut

First, we preprocess the RS image with principal component analysis (PCA) transform (Xu et al. 2014). This is followed by the automatic segmentation or manual tagging segmentation using the improved Grab-Cut mechanism. Finally, the detected edge can be saved to calculate area and circumference for convenience.

The novelty of our method lies first in the improvement of Grab-Cut mechanism. We propose to extend the ROI (region of interest) as background, which allows a considerably reduced degree of time consuming for a given quality of result. Second, we improve the Grab-Cut method to adapt the characteristics of large size and large amounts of information (e.g., Geographic Information) based on RS image. Finally, the improved Grab-Cut method is applied in detecting the marine sewage of RS image, which is competitive when compared with other state-of-the-art methods.

3 Improved Grab-Cut algorithm

3.1 Color data modeling

The RS image $z = (z_1, \dots, z_n, \dots, z_N)$ consists of pixels z_n in RGB color space. We use two Gaussian mixture models (GMMs) to model color data, one for the background and one for the foreground. To deal with the GMM conveniently, we introduce vector $k = \{k_1, \dots, k_n, \dots, k_N\}$ as each pixel's parameter, with $k_n \in \{1, \dots, K\}$.

The Gibbs energy function is as follows:

$$E(\alpha, k, \theta, z) = U(\alpha, k, \theta, z) + V(\alpha, z), \quad (1)$$

where $\alpha = (\alpha_1, \dots, \alpha_n, \dots, \alpha_N)$ is transparency and $\alpha_n \in \{0, 1\}$, with 0 for background and 1 for foreground. The data term U is defined as

$$U(\alpha, k, \theta, z) = \sum_n D(\alpha_n, k_n, \theta, z_n), \tag{2}$$

$$\begin{aligned} D(\alpha_n, k_n, \theta, z_n) &= -\log p(z_n | \alpha_n, k_n, \theta) - \log \pi(\alpha_n, k_n) \\ &= -\log \pi(\alpha_n, k_n) + \frac{1}{2} \log \det \Sigma(\alpha_n, k_n) \\ &\quad + \frac{1}{2} [z_n - \mu(\alpha_n, k_n)]' \Sigma(\alpha_n, k_n)^{-1} [z_n - \mu(\alpha_n, k_n)], \end{aligned} \tag{3}$$

where $p(\cdot)$ follows Gaussian distribution, and $\pi(\cdot)$ are mixture weighting coefficients.

The smoothness term can be written as

$$V(\alpha, z) = \gamma \sum_{(m,n) \in \mathbf{C}} [\alpha_n \neq \alpha_m] \exp(-\beta \|z_m - z_n\|^2), \tag{4}$$

where the constant $\gamma = 50$ is set by optimizing performance against ground truth over a training set. $\beta = (2(\|z_m - z_n\|^2))^{-1}$ and $\langle \cdot \rangle$ denotes the expectation over a remote-sensing image sample. \mathbf{C} is the set of pairs of neighboring pixels.

Therefore, the parameters of the model are given by

$$\theta = \{\pi(\alpha, k), \mu(\alpha, k), \Sigma(\alpha, k), \alpha \in \{0, 1\}, k = 1, \dots, K\}. \tag{5}$$

3.2 Energy minimization iteratively

Unlike Graph-Cut, Grab-Cut minimizes energy function iteratively. Therefore, the newly labeled pixels from the T_U region of the initial trimap will be used to modify the color GMM parameters.

3.3 User interaction and incomplete trimap

Incomplete labeling replacing complete trimap brings more flexibility. User only needs to define background T_B , leaving foreground $T_F = 0$. No hard foreground labeling is needed at all. Iterative energy minimization allows T_U representing foreground area T_F , and the labels of background area T_B are fixed. The initial value T_B is specified by user with a rough rectangle. If the initial information the user gives is not enough to get satisfactory results, the user needs to do more interactive job and provides more information.

3.4 Extension of ROI

According to the characteristics of RS image, ROI manually selected by the bounding box is insufficient. Large-scale images always have lower efficiency. For that matter, we propose to extend the ROI. Based on the characteristic of target, the size of original ROI will be doubled, less than the size of original image (e.g., the grid area in Fig. 1). Therefore, only the extension of ROI is involved in the calculation of the model. Experimental results show that the extension greatly improves efficiency of the method.

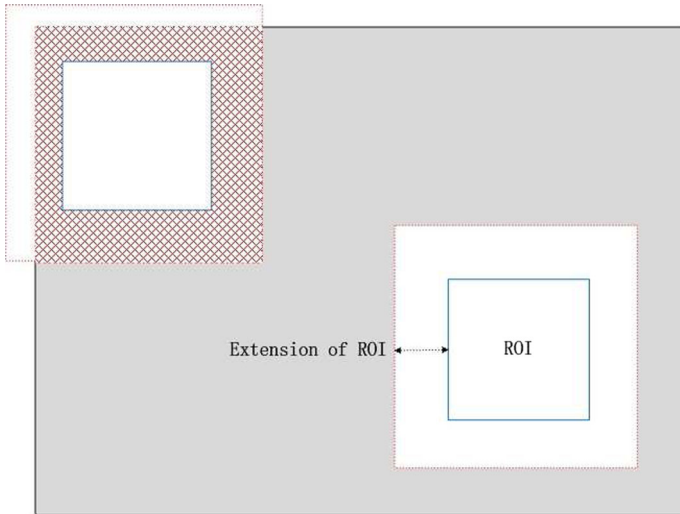


Fig. 1 Extension of ROI diagram

Table 1 Algorithm 1 Grab-Cut with the local extension of ROI

Input: RS image z , consists of pixels z_n in RGB color space

1: Manually select ROI \tilde{z} by the bounding box.

2: **while** not convergence **do**.

3: Assign GMM components to pixels: for each n in T_U

4: $k_n := \underset{k_n}{\operatorname{argmin}} D_n(\alpha_n, k_n, \theta, \tilde{z}_n)$.

5: Learn GMM parameters from data \tilde{z}

6: $\theta := \underset{\theta}{\operatorname{argmin}} U(\alpha, k, \theta, \tilde{z})$.

7: Estimate segmentation: use min cut to solve

8: $\min_{\{\alpha_n, n \in T_U\}} \min_k E(\alpha, k, \theta, \tilde{z})$.

9: **end while**

10: **Output:** α , the transparency

Table 1 shows the summary of the improved Grab-Cut:

4 Experiments

4.1 RS data

The Landsat-8 satellite of USA was launched on 11th February 2013, which carries a two-sensor payload, the Operational Land Imager (OLI), and the Thermal Infrared Sensor (TIRS). The reflectance of Landsat-8 OLI and TIRS was measured in 11 spectral bands: coastal/aerosol (0.44–0.45 μm), blue (0.45–0.51 μm), green (0.53–0.59 μm), red (0.64–0.67 μm), NIR (0.85–0.88 μm), SWIR (1.57–1.65 and 2.11–2.29 μm), cirrus (1.36–1.38 μm),

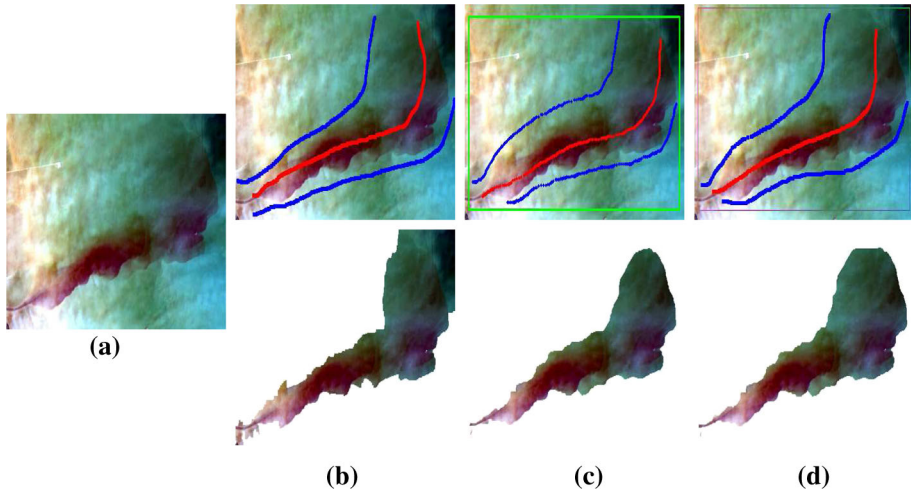


Fig. 2 Comparison of our method with various state-of-the-art methods. **a** Input RS image, **b** Boykov and Jolly's method, **c** Rother, Blake, and Kolmogorov's method, and **d** our method

thermal infrared (10.6–11.19 and 11.5–12.51 μm), and panchromatic mode (0.5–0.68 μm). In this study, blue band 2, green band 3, and red band 4 at 30 m spatial resolution of Landsat-8 image with cloud cover 6.56% were selected as the suitable data to detect the marine sewage. The image data were collected from the Bohai Bay near Tianjin city, China at 05:36:45 on August 9th, 2013.

4.2 Edge detection of the marine sewage

Figure 2 shows the comparison of our method with various state-of-the-art methods, including Boykov and Jolly's method and Rother, Blake, and Kolmogorov's method. There are still some visible errors (the upper right corner) in Boykov and Jolly's method. Rother, Blake, and Kolmogorov's method cannot detect the marine sewage in light color. Comparatively, our method is more accurate.

The experimental result on RS image with PCA transform can be seen in Fig. 3. The area of marine sewage is more obvious due to the PCA transform. It can be seen from Fig. 2b, c that Boykov and Jolly's method and Rother, Blake, and Kolmogorov's method cannot distinguish coast from marine sewage due to the similarity in color. Our method has better performances in this case, as shown in Fig. 2d.

As shown in Fig. 4, the area and circumference of marine sewage change with the numbers of iterations. Both of them tend to decrease with each iteration. As a result, the shape of marine sewage approaches the real results gradually.

4.3 The size of ROI extension

It can be concluded from Figs. 5 and 6 that our method may be incredibly inefficient for large-scale RS image. Time consuming increases as the size of ROI Extension becomes larger. The local adaptive extension of ROI just solved the problem. We choose the right size of ROI extension for our method (eg. doubled the ROI size) to improve efficiency.

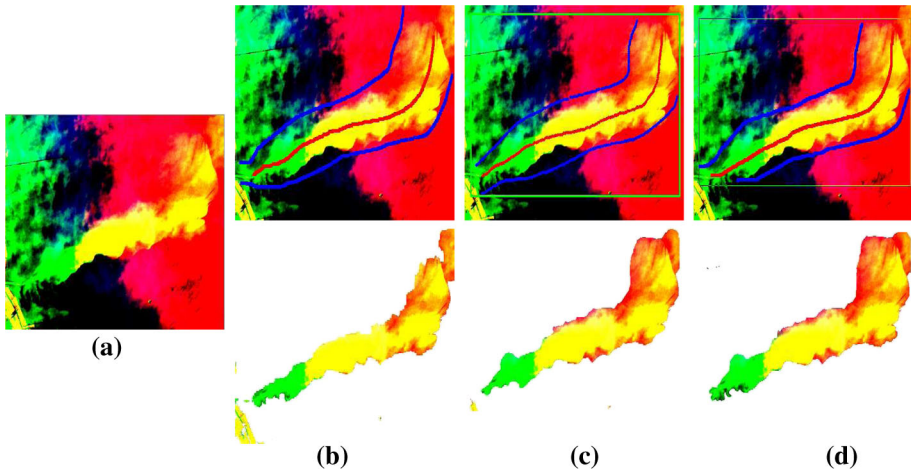


Fig. 3 Comparison of our method with various state-of-the-art methods. **a** RS image of PCA transform, **b** Boykov and Jolly's method, **c** Rother, Blake, and Kolmogorov's method and **d** our method

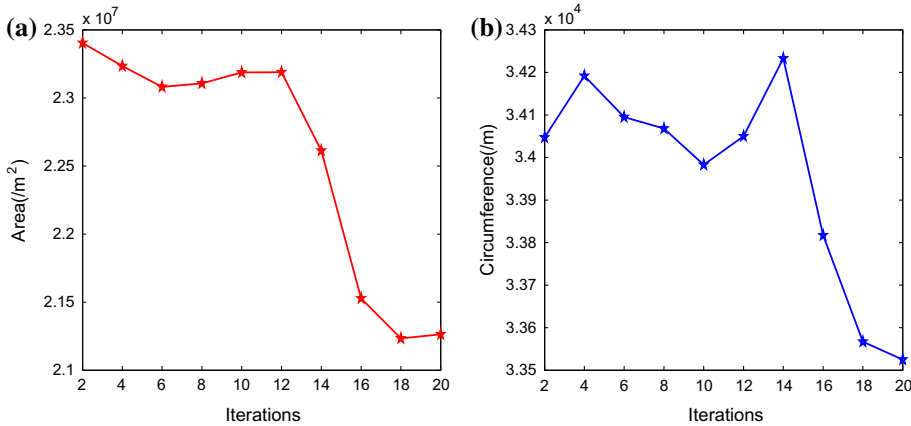
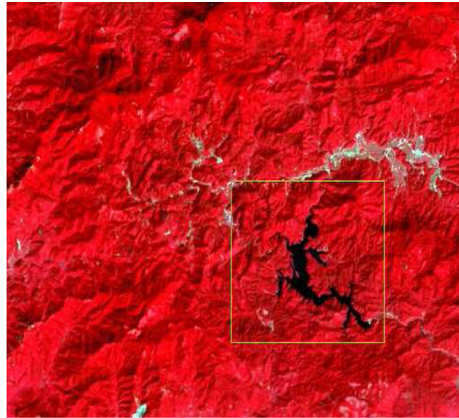


Fig. 4 Area and circumference of each iteration

4.4 Comparison with Grab-Cut method

The comparison of our method with Grab-Cut method on a river of Hainan Province in China is shown in Fig. 7a. In the first experiment, the time-consuming task of one segmentation without manually tagging was evaluated, as shown in Fig. 7b. As can be seen, the result of Grab-Cut method is not fine enough (i.e., the part of vegetation has not been removed). In contrast, our method can be used to accomplish this task. The river can be clearly distinguished. The second experiment was conducted with manually tagging which labels possible foreground and background. Grab-Cut still cannot remove the vegetation besides the river, as shown in Fig. 7c.

Given the time consuming of two experiments, our method costs 1–2% the time compared with that of Grab-Cut method. The improved Grab-Cut method can improve the efficiency significantly. The research code is implemented in C++ and tested under Windows environment with 3.20 GHz CPU and 4.00G RAM (Table 2).



(a) Input RS image

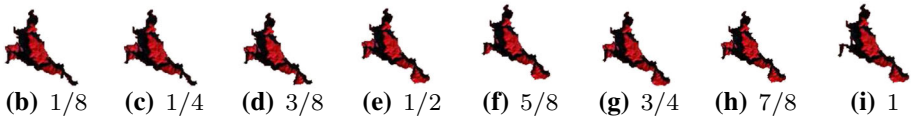
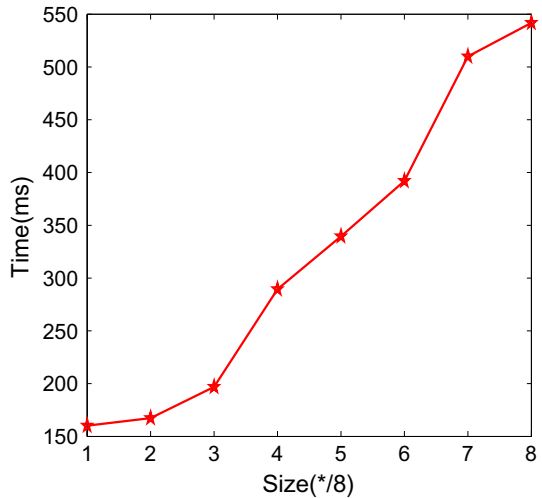


Fig. 5 Different sizes of ROI extension on a river of Hainan Province

Fig. 6 Time consuming of different sizes of ROI extension



5 Conclusion

In conclusion, a new and effective method for edge detection from RS image is proposed, which can be used to obtain foreground alpha mattes of good quality for large-scale images with a rather modest degree of user effort. Accordingly, software has also been developed base on the method that can be applied on the detection of marine sewage.

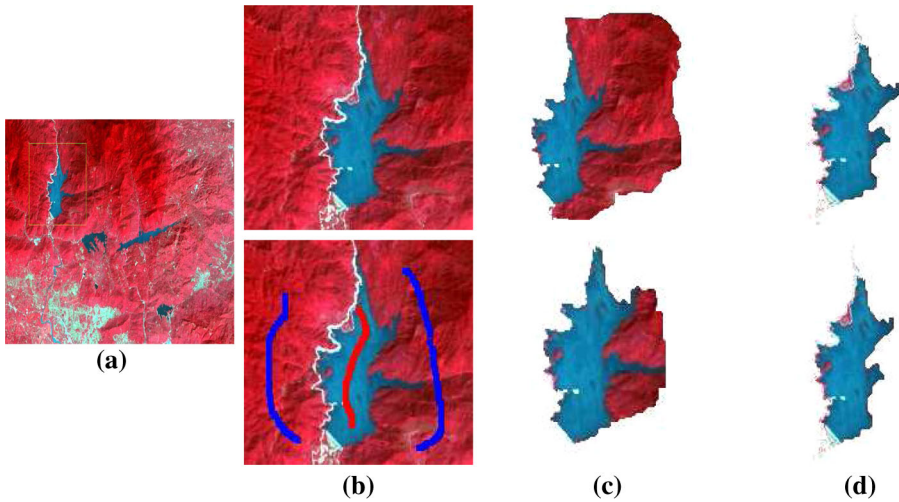


Fig. 7 Comparison of our method with Grab-Cut method on a river of Hainan Province in China. **a** Input RS image, **b** RS image with automatic segmentation and manual tagging, **c** Grab-Cut method, and **d** our method

Table 2 Comparison of our method with Grab-Cut method

Methods	Grab-Cut	Our method
Auto time (ms)	36,548.07	504.66
Labor time (ms)	25,020.83	358.26

Acknowledgements This work was supported by the National Natural Science Foundation of China under Grant No. 61379014 and the Natural Science Foundation of Tianjin under Grant No. 16JCYBJC15900. The authors express their sincere gratitude to Ying Li and Yumeng Song. Their very valuable suggestions and carefully reading helped the improvement of the paper. Further acknowledgement is extended to the reviews for their insightful comments which greatly improved the paper.

References

- Boykov YY, Jolly MP (1999) Multiway cut for stereo and motion with slanted surfaces. In: Proceedings of the seventh IEEE international conference on computer vision, vol 1. pp 489–495
- Boykov YY, Jolly MP (2001) Interactive graph cuts for optimal boundary & region segmentation of objects in ND images. In: Proceedings of eighth IEEE international conference on computer vision, vol 1. pp 105–112
- Boykov Y, Funka-Lea G (2006) Graph cuts and efficient ND image segmentation. *Int J Comput Vis* 70(2):109–131
- Boykov Y, Veksler O, Zabih R (2010) Fast approximate energy minimization via graph cuts. *IEEE Trans Pattern Anal Mach Intell* 23:1222–1239
- Chan TF, Vese LA (2001) Active contours without edges. *IEEE Trans Image Process* 10(2):266–277
- Grady L (2006) Random walks for image segmentation. In: *IEEE transactions on pattern analysis and machine intelligence*, vol 28, no 11. pp 1768–1783
- Juan O, Boykov Y (2006) Active graph cuts. In: *IEEE computer society conference on computer vision and pattern recognition*, vol 1, pp 1023–1029
- Kolmogorov V, Zabih R (2004) What energy functions can be minimized via graph cuts. *IEEE Trans Pattern Anal Mach Intell* 26:147–159

- Kolmogorov V, Boykov Y (2005) What metrics can be approximated by geo-cuts, or global optimization of length/area and flux. In: Tenth IEEE international conference on computer vision (ICCV), vol 1. pp 564–571
- Kolmogorov V, Boykov Y, Rother C (2007) Applications of parametric maxflow in computer vision. In: Proceedings of the 11th international conference on computer vision. Pittsburgh, Pennsylvania, USA. Proceedings. IEEE. pp 1–8
- Mumford D, Shah J (1989) Optimal approximations by piecewise smooth functions and associated variational problems. *Commun Pure Appl Math* 42(5):577–685
- Rother C, Kolmogorov V, Blake A (2004) Grabcut: interactive foreground extraction using iterated graph cuts. In: *ACM transactions on graphics (TOG)*, vol 23, no (3). pp 309–314
- Slabaugh G, Unal G (2005) Graph cuts segmentation using an elliptical shape prior. In: IEEE international conference on image processing (ICIP), vol 2. pp II–1222
- Unger M, Pock T, Trobin W et al (2008) TVSeg-Interactive Total Variation Based Image Segmentation. In: *Proceeding of BMVC*, vol 31. pp 44–46
- Xu J, Sun X, Zhang D et al (2014) Automatic detection of inshore ships in high-resolution remote sensing images using robust invariant generalized hough transform. *Geosci Remote Sens Lett* 11(12):2070–2074