

RESEARCH

Open Access

# Weakly supervised object extraction with iterative contour prior for remote sensing images

Chu He<sup>1,2\*</sup>, Yu Zhang<sup>1</sup>, Bo Shi<sup>1</sup>, Xin Su<sup>3</sup>, Xin Xu<sup>1</sup> and Mingsheng Liao<sup>2</sup>

## Abstract

This article presents a weakly supervised approach based on Markov random field model for the extraction of objects (e.g., aircrafts) in optical remote sensing images. This approach is capable of localizing and then segmenting objects in optical remote sensing images by relying only on several object samples without artificial labels. However, unlike direct combinations of object detection and segmentation, the proposed method develops a contour prior model based on detection results, thereby improving segmentation performance. Furthermore, we iteratively update the contour prior information based on the expectation-maximization algorithm. Numerical experiments illustrate that the proposed method can successfully be applied to the extraction of aircrafts in optical remote sensing images.

## 1 Introduction

Object detection and segmentation have received considerable attention as important procedures in automatic object identification in such fields as computer vision, remote sensing image processing, and so on. Based on the large number of works to which object detection and segmentation have been applied, a key distinction between these two methods can be found; object segmentation is usually interactive and incorporates guidance from the user throughout the analysis process, such as in GraphCut [1] and Snake [2], whereas object detection needs learning samples and/or supervising information from the user at the beginning of the analysis, such as in latent support vector machine LSVMs [3], Wu et al.'s Active Basis [4]. Nevertheless, object detection and object segmentation share numerous theoretical and methodological features, which if explored will be of benefit to each other. In this article, object detection results based on Active Basis [4] are developed to replace supervised learning samples in object segmentation. Object segmentation results can be obtained by providing several object samples. Furthermore, this combination employs a contour prior model

based on the detection results, thereby improving the segmentation performance.

From a methodological perspective, the main idea of numerous methods that have recently been used for object detection and segmentation can be divided into shape-based methods and feature-based methods. Shape-based methods, such as Felzenszwalb et al.'s LSVMs [3], Wu et al.'s Active Basis [4], Laptev et al.'s Snake [2], and Ferrari's kAS [5], exploit shapes similarities between objects by using different strategies and then obtain segmentation results through by connecting the segments. Shape-based methods are completed automatically without the need for human assistance. However, these methods hardly obtain segmentation results. Feature-based methods such as Cheng et al.'s [6] hierarchical lane detection system, Hassaballah et al.'s [7] independent components analysis, Borenstein and Ullman's [8] top-down bottom-up segmentation, Weisenssel et al.'s [9] Markov random field (MRF) model-based method, and Jia and Hong-qi's [10] interactive segmentation based on graph cuts, utilize different representations (color or texture features, distribution model) of image pixels or regions to distinguish objects from the background. However, these methods all require human assistance or strong supervised information.

Based on Active Basis [4], we developed detection results by using Morph-ActiveBasis [11] and thus

\*Correspondence: ChuHe@whu.edu.cn

<sup>1</sup>School of Electronic Information, Wuhan University, Wuhan 430079, P.R. China

<sup>2</sup>The State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

Full list of author information is available at the end of the article

proposed a contour prior model to improve segmentation performance, which will be detailed in Section 3. This article combines object detection (Morph-ActiveBasis [11]) and object segmentation (MRF) and proposes a contour prior model by using the above combination to improve segmentation performance.

## 2 Morph-ActiveBasis: from fragments to rough contours

Morph-ActiveBasis, which is presented in [11], is based on Active Basis. Morph-ActiveBasis can determine the basic edge contours of objects from a set of object samples without the need for artificial labels and can detect similar objects in given images. Unlike the scattered fragments obtained by using the original Active Basis, Morph-ActiveBasis employs fragment connection to link scattered fragments thus forming a sketch of the object contour (for details, see [11]).

### 2.1 Fragment detection

The Active Basis [4] model is utilized to detect the basic edge contours of objects. Active Basis represents contours through a set of Gabor Wavelet bases [12,13]. Moreover, Active Basis does not require human guidance and can automatically detect objects in the image. However, the detected results are only scattered fragments that could not represent the integral contour of object. Therefore, we propose the use of fragment connection to link the scattered fragments, thus forming a sketch of the object contour.

### 2.2 Fragment connection

The principle of fragment connection in [11] is based on the structure information among fragments.

### 2.3 Rough contours extraction

The Morph-ActiveBasis detection algorithm estimates the Gaussian distribution models of objects and backgrounds

by using contour sketches, and then segments the contours by utilizing GraphCut segmentation [14] algorithm. Given an image  $I = \{s_1, s_2, \dots, s_{N_I}\}$ ,  $s_i$  is the  $i$ th pixel in the image. The segmentation result is that each one of  $N_I$  pixels is assigned to a label  $y_i \in \{0, 1\}$ , where 0 and 1 represent objects and backgrounds, respectively, thus the segmentation yields result  $Y = \{y_1, y_2, \dots, y_{N_I}\}$ .

$$\bar{Y} = \arg \max_Y \{P(Y|I, \Theta)\} \quad (1)$$

$$P(Y|I, \Theta) \propto \prod_{i=1}^{N_I} p_{L_i}(y_i|s_i, \Theta) p_{N_I}(y_i|y_{N_I}) \quad (2)$$

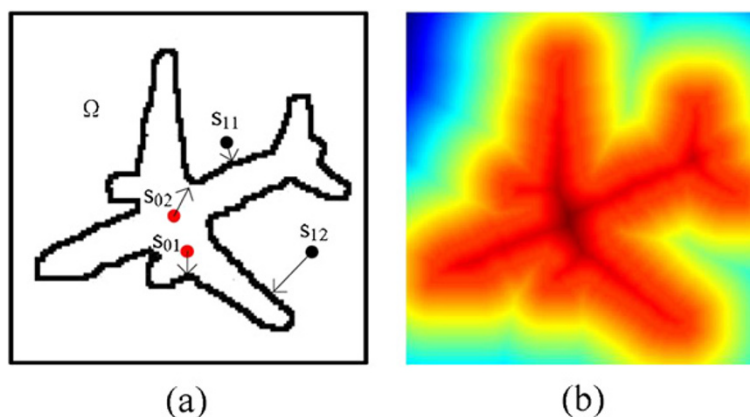
Here,  $\Theta$  is the parameter of the Gaussian distribution  $\Theta = \{\mu, \sigma\}$ , and  $\mu, \sigma$ , respectively, represent the mean and variance of Gaussian distribution. The likelihood probability  $p_{L_i}$  can be obtained by using Gaussian distribution model of objects and background, as shown in Equation (6).  $p_{N_I}$  denotes the Potts model.

## 3 Expectation-maximization (EM) contour MRF: from rough contours to further segmentation

In this section, we present a contour prior model of objects to improve segmentation performance under the MRF model framework. The idea of this contour prior model is to assign pixels (that are located inside an object contour) with a higher probability to become *object*. By contrast, pixels outside the contour are assigned a higher probability to be *background*. The probability of the contour prior is based on the distance between each pixel and its nearest contour point.

### 3.1 Contour prior information based on rough contours

As shown in Figure 1a,  $\Omega$  is the initial contour of the object,  $s_{01}, s_{02}, s_{11}, s_{12}$  are the four arbitrary pixels in the image ( $s_{01}, s_{02}$  are located inside  $\Omega$ , whereas  $s_{11}, s_{12}$  are outside  $\Omega$ ) and  $d_{01}, d_{02}, d_{11}, d_{12}$  denote the distances from



**Figure 1 Sketch map of contour prior. (a)** Distance from point to contour; **(b)** probability map of shape prior.

$s_{01}, s_{02}, s_{11}, s_{12}$  to  $\Omega$ , respectively. Inside the contour, a pixel that is located farther from the contour is more likely to become an object. However, outside the contour, a pixel that is farther from the contour has higher probability of becoming a background. For instance, in Figure 1a,  $s_{02}$  is more likely to become an object as compared with  $s_{01}$  because  $d_{01} < d_{02}$ , whereas  $s_{12}$  is more likely to become a background as compared with  $s_{11}$  because  $d_{11} < d_{12}$ .

Figure 1b shows the sketch map of the shape prior probability  $p_{s_i}$ . The data shown in Figure 1 can mathematically be expressed by Equations (3)–(5) such as in [15,16]. Prior knowledge on the background and the object is shown in Equations (3) and (4), respectively. Equation (5) gives the distance of the pixel to the contour, where  $\mu$  is a constant that indicates the distance coefficient  $sign(s_i, \Omega) = 1$  when  $s_i$  is inside  $\Omega$  and  $sign(s_i, \Omega) = -1$  when  $s_i$  is outside  $\Omega$ ; and  $loc(s)$  is the coordinate position of pixel  $s$  in the image.

$$p_{s_i}(y_i = 0|\Omega) \propto \frac{1}{1 + \exp(\mu * \text{dist}(s_i, \Omega))} \quad (3)$$

$$p_{s_i}(y_i = 1|\Omega) \propto 1 - \frac{1}{1 + \exp(\mu * \text{dist}(s_i, \Omega))} \quad (4)$$

$$\text{dist}(s_i, \Omega) = \text{sign}(s_i, \Omega) \min_{s \in \Omega} \|\text{loc}(s_i) - \text{loc}(s)\| \quad (5)$$

### 3.2 Segmentation based on the MRF model and on EM iteration

To obtain better segmentation results from the contour prior information, we combine the EM algorithm [17] with GraphCut optimization. We apply the segmentation result to update the parameters of the Gaussian distribution model and then utilize the new distribution parameters to initialize GraphCut optimization. The E- and M-steps are as follows:

**E-step:** The likelihood probability of pixels to class  $y_i$  is computed based on the current distribution parameters of Gaussian distribution  $\Theta^{(t)} = \{\mu^{(t)}, \sigma^{(t)}\}$ , where  $\mu, \sigma$ , respectively, represent the mean and variance of Gaussian distribution and express as follows:

$$p_{L_i}(y_i|s_i, \Theta^{(t)}) \propto \frac{p(s_i|y_i, \Theta^{(t)})}{\sum_{c=0}^1 p(s_i|y_i = c, \Theta^{(t)})} \quad (6)$$

**M-step:** Under the framework of the MRF model, we denote each pixel as a node in the MRF model and describe its likelihood probability with  $p_{L_i}$ . The relationship between pixels (nodes) is represented by the combination of  $p_{s_i}$  and  $p_{N_i}$ , which is a pairwise probability in MRF. The probability  $p_{L_i}$  which is obtained in the E-step is utilized to combined with the contour prior  $p_{s_i}$  of the object and the prior probability  $p_{N_i}$  of the Potts model to obtain the MRF model in Equations (7) and (8). The GraphCut optimization algorithm is then applied to solve Equations (7) and (8), thus obtaining the new labeling data  $Y$ . The new Gaussian distribution parameters  $\Theta^{(t+1)} = \{\mu^{(t+1)}, \sigma^{(t+1)}\}$  of objects and background are computed,

and the shape  $\Omega^{(t+1)}$  of the object is updated. The EM algorithm continuously repeats the E- and the M-steps until the convergence condition is satisfied.

$$\bar{Y} = \arg \max_Y \{P(Y|I)P(I/\Omega^{(t-1)})P(Y)\} \quad (7)$$

$$P(Y^{(t)}|I, \Omega^{t-1}, \Theta^{t-1}) = \exp \left[ \sum_{i \in I} \left[ \alpha f(s_i|y_i, \Theta^{t-1}) + \beta f(y_i|d(s_i, \Omega^{t-1})) + \gamma \sum_{j \in C_i} \delta(y_i, y_j) \right] \right] \quad (8)$$

The right-hand side of Equation (8) is consisted of three parts which, respectively, represent  $P_{L_i}$ ,  $P_{s_i}$ , and  $P_{N_i}$ . The first two parts, respectively, represent Gaussian probability and contour, the last one is the Potts prior model.  $C_i$  is the cliques of pixel  $s_i$ . We use eight neighborhood cliques. Then we defined the potential or energy function for these cliques through Gibbs distribution which is computed by Potts prior model.  $d(s_i, \Omega^{t-1})$  is the distance from  $s_i$  to the contour  $\Omega^{t-1}$ ,  $\alpha, \beta, \gamma$  are the normalized parameters.

### 4 Flow of the proposed algorithm

The proposed weakly supervised object extraction algorithm aims at achieving improved object fragment detection and segmentation results with minimal supervision. As shown in Figure 2, the algorithm first utilizes fragment connection and GraphCut segmentation to obtain the initial contour  $\Omega^0$  of the object based on the Active Basis model. The algorithm then constructs the MRF on the initial contour of the object and adopts an iterative optimization approach combined with the EM algorithm to achieve improved object segmentation results. The algorithm process is shown as below:

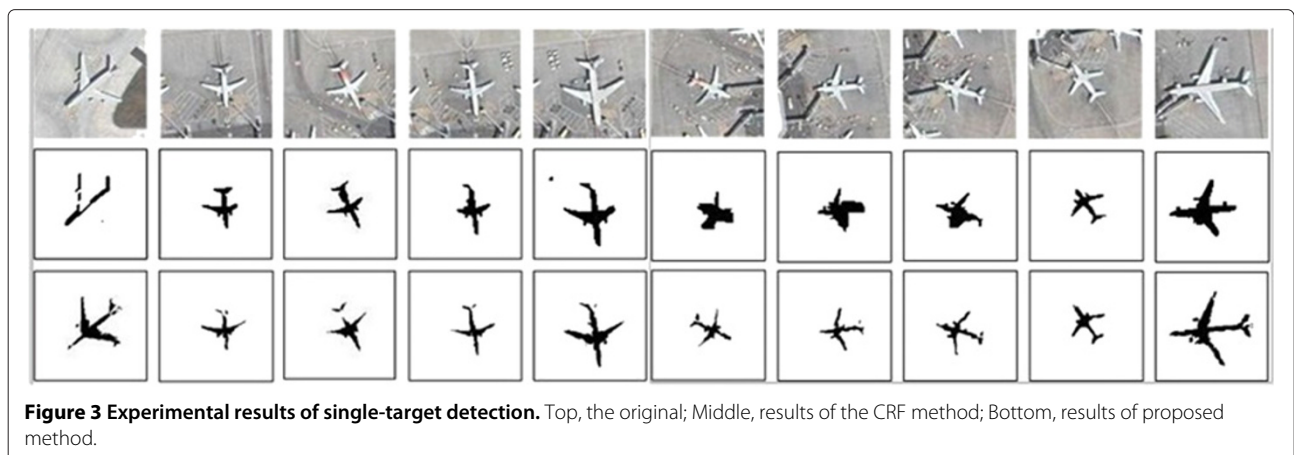
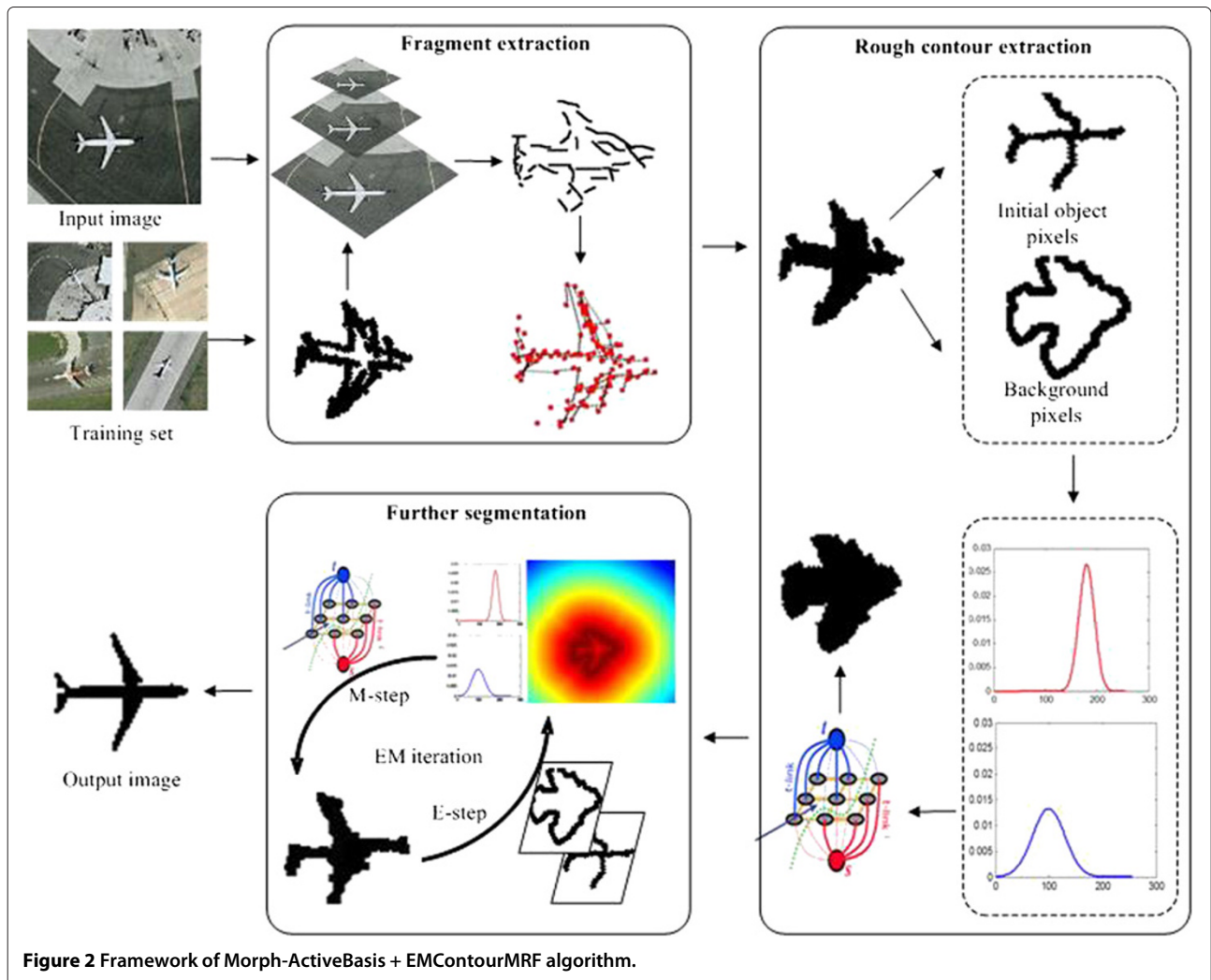
**Initialization:** Use Morph-ActiveBasis method to obtain the initial contour  $\Omega^0$  of objects

#### Step 1: E-Step:

- a: Use (3)–(5) to compute for the shape prior  $p_{s_i}$  of each pixel in the testing image based on the current contour  $\Omega^t$  of the object;
- b: Compute the likelihood  $p(s_i|y_i, \Theta)$  and posterior probability  $p_{L_i}$  of pixels according to the current distribution parameters of the objects and background, as shown in Equation (6);

#### Step 2: M-Step:

- a: Use of GraphCut optimization algorithm to solve (7)–(8) to obtain the new labeling data based on (3)–(6);



**Table 1 Completeness and correctness of object segmentation in the images taken from the Beijing Capital International Airport with a single aircraft**

Methods	Completeness (%)	Correctness (%)
CRF	72.14	88.29
Morph-ActiveBasis + EMContourMRF	72.10	94.08

b: Update object's shape  $\Omega^{(t+1)}$  as well as Gaussian distribution parameters  $\Theta^{(t+1)}$ ;

**Termination:** Repeat 2 and 3 until the change in the segmented object regions is less than a certain threshold, that is, the iterative convergence.  
**end**

## 5 Experimental results and analysis

### 5.1 Experimental data

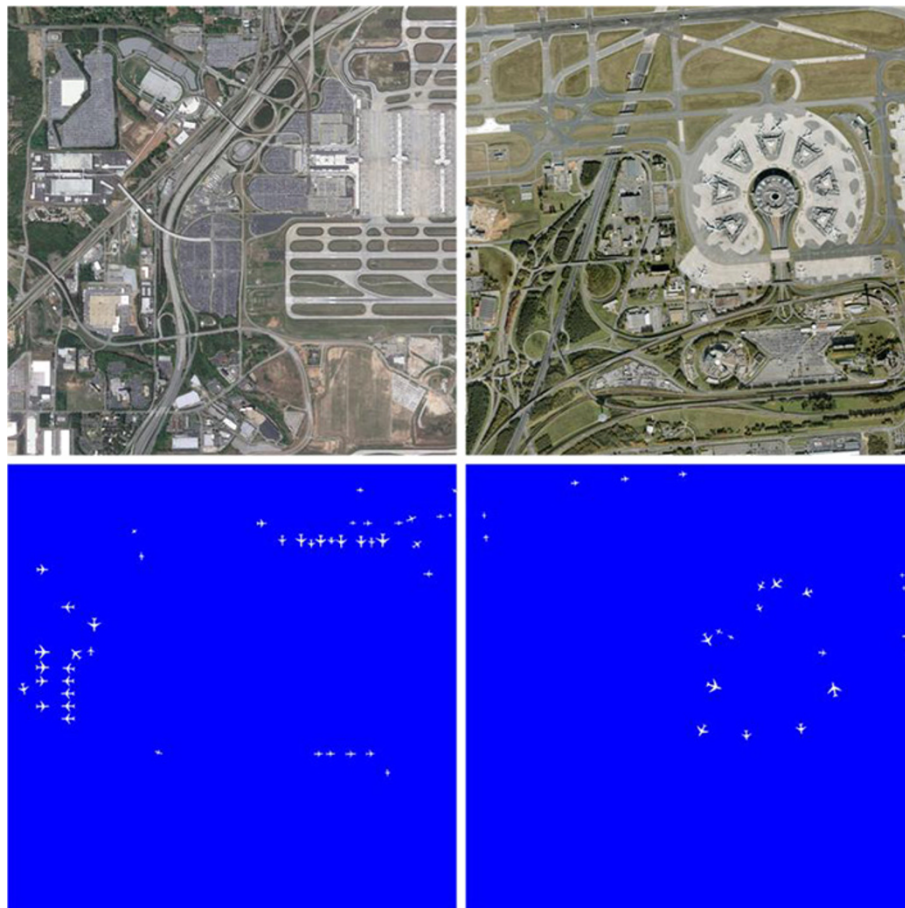
During the experiment, two new and significantly extended databases have been formed through capturing

screen shot images from Google earth. One database contains 100 small images with the size  $128 \times 128$  obtained from the Beijing Capital International Airport; the other database consists of 10 big images with the size  $5000 \times 5000$  obtained from Atlanta and other airports in the world. To obtain clearer images of the aircrafts, eye altitude is set to lower than 500 m when using Google Earth to obtain screenshots. Aircraft targets from the airport are selected as experimental targets from which the training and testing image samples are taken. At least one image is needed for training. We use ten images for training in the experiments.

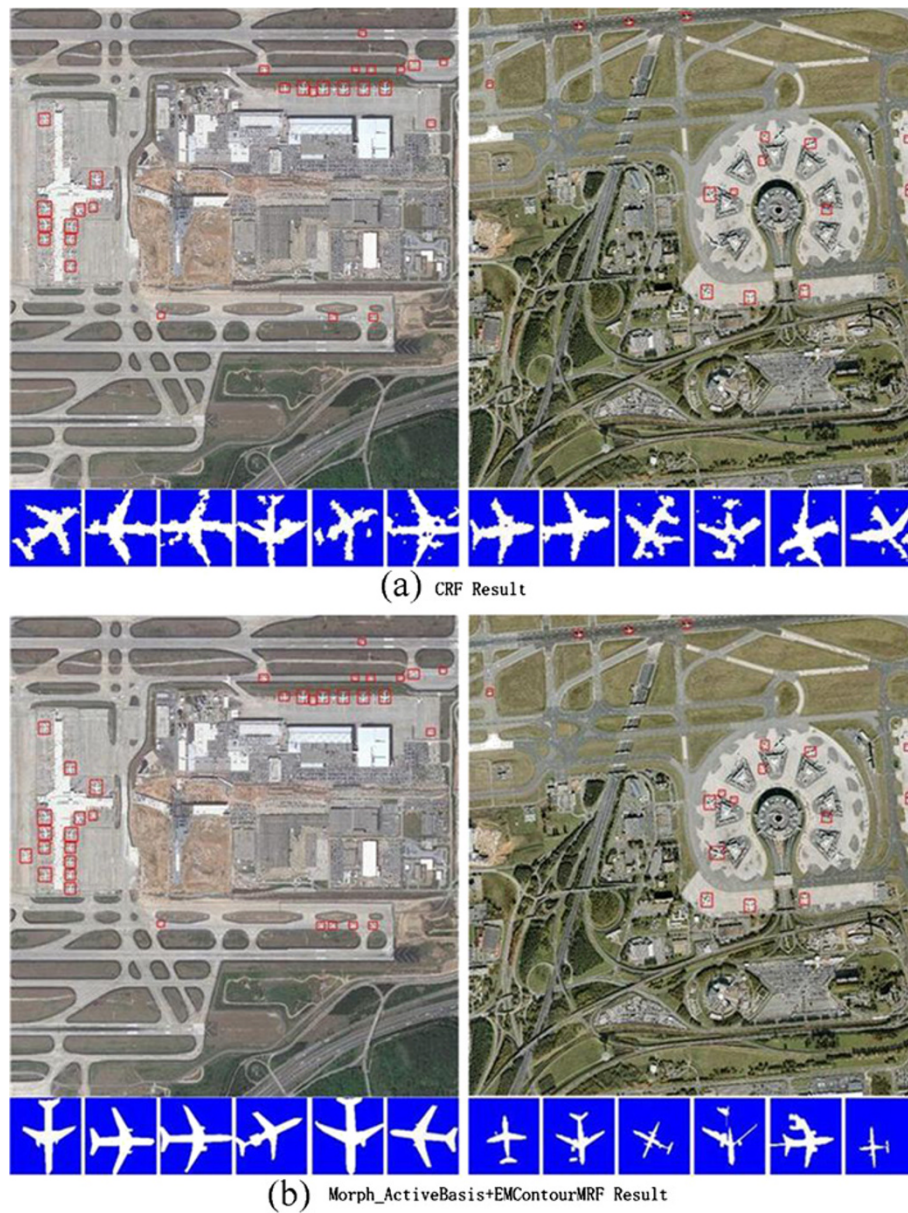
### 5.2 Experiment setting

Comparisons are performed by using the conditional random field (CRF) [18] with RCC prior. CRF needs artificial labels to start segmentation. For the same experimental setting, we used the initial object pixels and background pixels (in Step 2, as shown in Figure 2) as the input of CRF.

To evaluate quantitatively the performance of proposed method in terms of object segmentation, measures of



**Figure 4** Experimental images for multi-target detection. Top, Atlanta Airport and Charles de Gaulle Airport; Bottom, Ground Truth.



**Figure 5** Extraction results and some segmentation results of multi-aircrafts: (a) CRF approach and (b) Morph-ActiveBasis + EMContourMRF algorithm.

**Table 2** Detection rate and error detection rate of aircraft detection in the Atlanta Airport and Charles de Gaulle Airport

Image	Detection rate (%)	Error detection rate
Morph-ActiveBasis	81.54	0
CRF	61.15	0

**Table 3** Completeness and correctness of aircraft detection in the Atlanta Airport and Charles de Gaulle Airport

Methods	Completeness (%)	Correctness (%)
CRF	81.17	76.3
Morph-ActiveBasis + EMContourMRF	74.23	90.18

**Table 4 Completeness and correctness of aircraft detection in ten international airports**

Methods	Completeness (%)	Correctness (%)
CRF	90.26	72.91
Morph-ActiveBasis + EMContourMRF	73.16	85.57

completeness and correctness are adopted which are similar to recall and precision in image retrieval [19], in which completeness denotes the ratio of correctly segmented object pixels to the sum of the true object pixels, whereas correctness denotes the ratio of correctly segmented object pixels to the sum of pixels in the object segmentation results. True object pixels are manually labeled in the original image. The detection rate and error detection rate are also employed to assess target detection performance, where the former indicates the ratio of the number of detected target pixels to the total number of targets whereas the latter stands for the number of wrongly detected target pixels to the total number of targets.

### 5.3 Results

#### 5.3.1 Experiment 1

We first test the proposed method on 100 small images taken from Beijing International Airport (size:  $128 \times 128$ , each contains a single aircraft). Figure 3 shows ten samples of the experimental results. As shown in Figure 3, contours are relatively well segmented. The proposed Morph-ActiveBasis + EMContourMRF algorithm is compared with the CRF method. Table 1 shows the segmentation accuracy. The completeness of Morph-ActiveBasis is higher than that of CRF, but its correctness is significantly lower. Based on Morph-ActiveBasis, EM Contour MRF not only enhances the contour of objects, but also significantly improves the segmentation accuracy.

#### 5.3.2 Experiment 2

The second experiment involved the extraction of multiple targets from images taken from Atlanta Airport (size:  $5000 \times 5000$ , each contains more than one aircraft) as shown in Figure 4. Figure 5 displays the results. More aircrafts have correctly been detected by using Morph-ActiveBasis method than CRF approach. Table 2 shows the numbered bounding box, detection rate, and the error detection rate. Figure 5 illustrates some of the corresponding segmented results. The completeness and correctness of segmentation of each image are shown in Table 3, and the results of all the ten images are shown in Table 4.

Although the completeness of the propose method is lower than that of CRE, the former significantly improves correctness. We can learn that the weakly supervised

MRF model with iterative shape prior can also achieve satisfactory performance in multi-target extraction.

## 6 Conclusion

An automatic object extraction algorithm for optical remote sensing images is proposed in this article. First, the Active Basis algorithm was used to detect the fragment of object contour and to obtain the initial object contour by using fragment connection based on morphology. The Gaussian mixture models of objects and background were then built under the Bayesian framework and the prior information on object shapes was introduced. Finally, EM iteration and GraphCut optimization were combined for object segmentation. Our algorithm has the following advantages: (1) Morph-ActiveBasis needs only multiple images containing the objects or simple hand-drawn sketches of objects without other prior instructions. Less human intervention is required by the proposed method compared with other object segmentation methods that are based on feature classification; (2) Since EMContourMRF overcomes the problem of unorganized segmentation results, which occurs in general models that have only neighborhood prior, because EMContourMRF introduces prior information on object shapes and uses contours to constrain the segmentation process. Thus, EMContourMRF obtains better contours of objects; (3) EMContourMRF combines EM iteration and GraphCut optimization to estimate and to optimize the distribution models of objects and backgrounds repeatedly. EMContourMRF has relatively higher segmentation precision than object detection methods that are based on shape fragments.

The accuracy of object detection and segmentation is significantly affected by shadows. Thus, the removal of shadows is the focus of our follow-up work. Moreover, with the development of network information, automatic acquisition of images that contain more objects to construct a complete training database is a key step in our follow-up work to achieve better segmentation results independently.

#### Competing interests

The authors declare that they have no competing interests.

#### Acknowledgements

This study was supported by the National Basic Research Program of China (973 Program) under Grant No. 2013CB733404, NSFC grant (Nos. 60702041, 41174120, 41021061), the China Postdoctoral Science Foundation funded project, and the LIESMARS Special Research Funding.

#### Author details

<sup>1</sup>School of Electronic Information, Wuhan University, Wuhan 430079, P.R. China. <sup>2</sup>The State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China. <sup>3</sup>Institut Telecom, Telecom ParisTech, LTCI, Paris, France.

Received: 28 December 2011 Accepted: 5 November 2012  
Published: 13 February 2013

## References

1. C Rother, V Kolmogorov, AM Blake, Grabcut—interactive foreground extraction using iterated graph cuts. in *Proceedings of the ACM Siggraph*. **23**(3), 309–314 (2004)
2. I Laptev, H Mayer, T Lindeberg, W Eckstein, C Steger, A Baumgartner, Automatic extraction of roads from aerial images based on scale space and snakes. *Mach. Vis. Appl.* **12**(12), 23–31 (2000)
3. P Felzenszwalb, D McAllester, D Ramanan, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. A discriminatively trained, multiscale, deformable part model, (Anchorage, AK, 2008), pp. 1–8
4. YN Wu, ZZ Si, HF Gong, SC Zhu, Learning active basis model for object detection and recognition. *Int. J. Comput. Vis.* **90**(2), 198–235 (2010)
5. V Ferrari, L Fevrier, F Jurie, C Schmid, Groups of adjacent contour segments for object detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(1), 36–51 (2008)
6. HY Cheng, CC Yu, CC Tseng, KC Fan, JN Hwang, BS Jeng, Environment classification and hierarchical lane detection for structured and unstructured roads. *Inst. Eng. Technol.* **4**(1), 37–49 (2010)
7. M Hassaballah, T Kanazawa, S Ido, Efficient eye detection method based on grey intensity variance and independent components analysis. *Inst. Eng. Technol.* **4**(4), 261–271 (2010)
8. E Borenstein, S Ullman, in *European Conference on Computer Vision*, vol. 3. Learning to segment, (Prague, 2004), pp. 315–328
9. RA Weisenssel, WC Karl, DA Castanon, in *Algorithms for Synthetic Aperture Radar Imagery*, vol. 5. Gregory J. Power, Phil Douville, Markov random field segmentation methods for SAR target chips, (Orlando, FL, 1999), pp. 462–473
10. L Jia, W Hong-q, A graph cuts based interactive image segmentation method. *J. Electron. Inf. Technol.* **30**(8), 1973–1976 (2008)
11. C Qian, C He, XP Deng, H Sun, in *Multispectral Image Processing and Pattern Recognition*, vol. 3495. Object contour detection in remote sensing image (Yichang, China, 2009), pp. 1–6
12. TS Lee, Image representation using 2D Gabor wavelets. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**(10), 959–971 (1996)
13. M Lyons, S Akamatsu, M Kamachi, J Gyoba, in *Proceedings of the Conference on Automatic Face and Gesture Recognition*. Coding facial expressions with Gabor wavelets, (Nara, 1998), pp. 200–205
14. Y Boykov, O Veksler, R Zabih, Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(11), 1222–1239 (2001)
15. MP Kumar, PHS Torr, A Zisserman, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1. OBJ CUT (San Diego, CA, USA, 2005), pp. 18–25
16. MP Kumar, PHS Torr, A Zisserman, OBJ CUT: efficient segmentation using top-down and bottom-up cues. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**, 530–545 (2009)
17. A Dempster, N Laird, D Rubin, Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B.* **39**(1), 1–38 (1997)
18. J Inglda, J Michel, Qualitative spatial reasoning for high-resolution remote sensing analysis. *IEEE Trans. Geosci. Remote Sens.* **47**(2), 599–612 (2009)
19. LZ Lu, *Study on Remote Sensing Image-base Content Retrieval Based on Database Model*. (China Meteorological Press, Beijing, 2005), pp. 35–40

doi:10.1186/1687-6180-2013-19

**Cite this article as:** He et al.: Weakly supervised object extraction with iterative contour prior for remote sensing images. *EURASIP Journal on Advances in Signal Processing* 2013 **2013**:19.

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)