INTEGRATING SPARSE RECONSTRUCTION SALIENCY AND TARGET-AWARE ACTIVE CONTOUR MODEL FOR AIRPORT EXTRACTION

Qijian Zhang, Wenqi Shi, Libao Zhang, Yue Liu*

College of Information Science and Technology, Beijing Normal University, Beijing, 100875, China

ABSTRACT

This paper deals with automatic airport extraction in remote sensing images (RSIs). We present an innovative framework using sparse reconstruction saliency (SRS) and target-aware active contour model (TAACM). We begin with segmenting an image into superpixels and extracting the feature vectors. In feature space, we learn an airport target dictionary and a background dictionary for sparse representation of all image sub-regions. The saliency confidence can be determined by sparse reconstruction error. Based on the saliency map, we apply a novel target-aware active contour model (TAACM) for target contour tracking and provide accurate descriptions about the airport details. Extensive experiments demonstrate that the SRS algorithm outperforms nine competing saliency models in remote sensing scenes. In addition, the proposed airport extraction framework achieves higher detection accuracy compared with three competing methods.

Index Terms —Object extraction, saliency, dictionary learning, sparse coding, active contour model

1. INTRODUCTION

Automatic airport extraction in RSIs is receiving great attention in military and civil applications. To the best of our knowledge, traditional works can be classified into two categories: works based on edge or line segment detection [1, 2], and works based on scene segmentation [3]-[9].

Based on the observation that airport region of interest (ROI) consists of long, straight runways with continuous edges, the first category detects line segments [1] from the whole image and identify target area by means of a classifier such as support vector machine (SVM) [2]. This scheme achieves simple and fast detection, but is very sensitive to disturbing linear ground objects (e.g. residential areas, long rivers, and roads) in wide-range RSIs. In contrary, scene segmentation based methods utilize texture information [3] to obtain candidate sub-regions, from which the airport ROI is refined. In method [4] and [5], bottom-up and top-down saliency cues are combined to generate fusion saliency map, providing more efficient estimation of the airport fields. This scheme considers more contextual information and achieves better robustness and accuracy, but it becomes time-consuming as the data volume significantly increases.

Since the most discriminative feature representations of the airport targets are difficult to artificially predict and construct, deep learning based architecture [10, 11] received more and more attention recently. With adequate training samples and a well-designed network, deep learning method is able to excavate implicit target properties and is robust to very complicated image context. However, since there is no existing large open dataset targeted at airport in RSIs, the ability of deep learning model is often restricted by limited amount of training samples. In [11], the line group weighted saliency map is generated first, then deep residual learning network (ResNet) is employed to determine the airport from candidate regions. Although the ResNet serves as a better classifier, the overall effect of this work still highly relies on the saliency detection procedure.

Conclusively, many valuable researches on the topic of airport detection have been conducted. But we noticed some shortcomings of traditional works that are worth enhancing. First, many saliency-based detection methods employ the saliency models targeted at natural scene image, which is not fully suitable for RSIs. Second, most of existing airport detection frameworks only focus on enhancing the detection rate, failing to further extract detailed information of airport contours and structures.

To address the problems above, this paper presents a novel airport extraction framework via sparse reconstruction saliency (SRS) model and target-aware active contour model (TAACM). In the first step, the remote sensing image is segmented into superpixels. Each superpixel is represented by a feature vector in terms of color, contrast, and gradient magnitude metrics. The target dictionary and background dictionary are learned from the feature matrix respectively for the sparse representation of all image sub-regions. In the second step, we compute the sparse coding coefficients for each region and assign saliency confidence by the sparse reconstruction error. The saliency map is supposed to roughly locate the airport field and suppress background noises. Finally, we introduce saliency priors into the traditional energy function of active contour model and construct a novel target-aware active contour model (TAACM) for target contour tracking. By energy minimization, the contour curves can be pulled towards the airport so that we can separate highly-accurate airport structures from candidate ROIs.

*Corresponding author: libaozhang@163.com. This research was supported by Beijing Natural Science Foundation (4162033) and National Natural Science Foundation of China (41771407 and 61571050).

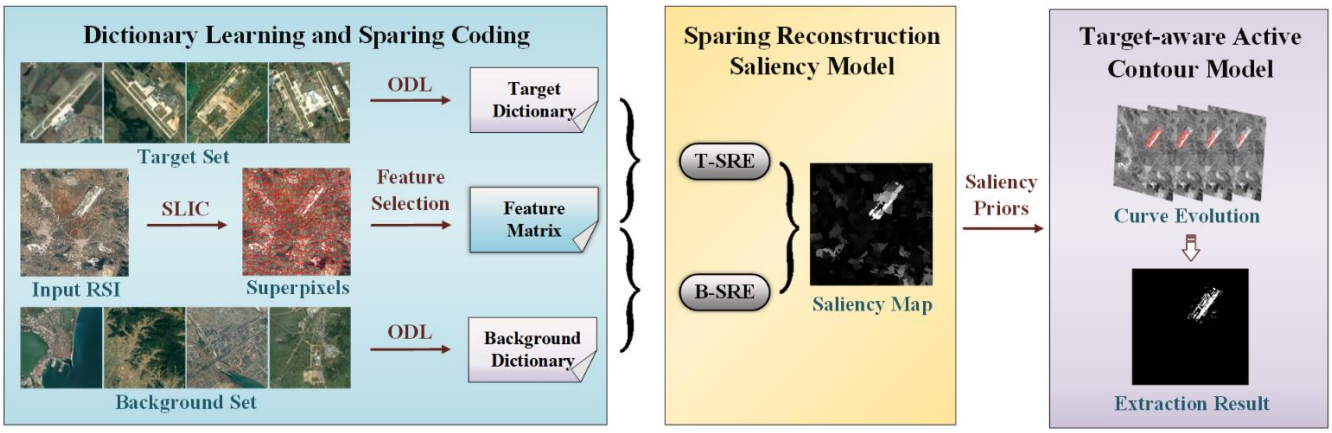


Fig. 1. *Flowchart of the proposed airport extraction framework.*

2. METHODOLOGY

As illustrated in Fig. 1, we start by segmenting the whole image context into superpixels by SLIC algorithm [12], a widely applied method that generates regular and concise superpixels with well-preserved boundaries. Note that in our implementation, a bilateral filter is employed to pre-process the input RSI. This method helps to suppress low-frequency noises and preserve edge elements, avoiding degrading the airport runways structures.

We design a set of feature selection metrics to represent the superpixel with a high-dimension feature vector. In the feature space, pre-learned target dictionary and background dictionary are utilized to compute the sparse representation coefficients for all image sub-regions. By measuring the target sparse reconstruction error (T-SRE) and background sparse reconstruction error (B-SRE), we can assign saliency confidence to each superpixel and generate saliency map to estimate the airport location. Finally, we employ the proposed TAACM to acquire the airport contours where saliency priors are introduced to guide the evolving curves to move towards the airport target.

2.1. Target and Background Dictionary Learning

In the dictionary learning stage, we build up the target and background training set respectively by collecting airport and background RSI patches. In order to efficiently express and distinguish the characteristics of different types of the patches, we design a feature representation pattern to extract the feature vector from the superpixel.

Suppose that we obtain N superpixels from the input as $\Gamma = \{s_1, s_2, ..., s_N\}$. For superpixel s_i, we integrate color feature, global contrast feature, and gradient magnitude feature to construct the feature vector x_i as

$$x_i = \{c^{lab}, \zeta^c, \chi^m\} \tag{1}$$

where $c^{lab} = (c^l, c^a, c^b)$ denotes the mean color feature vector of s_i in *CIELAB* color space. ζ^c measures the global contrast level. We collect the color information of all pixels in s_i and construct the color histogram. Then we quantize each color channel into 12 discrete values as a 36-d feature vector h_i. Let ϖ_i be the total number of pixels in s_i, the ζ^c can be formulated as

$$\zeta^c(s_i) = \sum_{s_k \in \Gamma} \varpi_i \left\| h_i, h_k \right\|_2 \tag{2}$$

Moreover, χ^m denotes the mean gradient magnitude obtained by Gabor filtering. Conclusively, the formulation of x_i implements the observation that the airport area has relatively uniform and distinct color distribution, higher contrast to the whole image context, and strong gradient magnitude. Thus, the superpixel set can be represented as feature matrix $\Gamma^f = \{x_1, x_2, ..., x_N\}$.

After that, we employ an efficient dictionary learning algorithm, online dictionary learning (ODL) presented in work [13], to learn the target dictionary D^t and background dictionary D^b from the corresponding training set. And the pre-trained dictionaries are supposed to provide appropriate bases for effective sparse coding of the airport ROI and the remote sensing background.

2.2. Sparse Reconstruction Saliency (SRS)

Sparse representation aims to describe the original signal with the least amount of basis atoms under a given over-complete dictionary. Inspired by the algorithm in [14] that determines saliency via boundary template based sparse reconstruction error (SRE), this paper simultaneously computes the T-SRE and B-SRE with respect to the given target and background dictionary, and generates saliency map by the two metrics.

We notice that the target superpixels should be better expressed by the target dictionary with less reconstruction error than that with the background dictionary. Similarly,

the background objects can be reconstructed by the background dictionary more accurately. Hence, salient superpixels will correspond to lower T-SRE and higher B-SRE.

Based on the knowledge above, we first acquire the target and background sparse representation coefficients, α_i^t and α_i^b, for the i^{th} superpixel as

$$\alpha_i^{t/b} = \arg\min_{\alpha_i^{t/b}} \left\| x_i - D^{t/b}\alpha_i^{t/b} \right\|_2^2 + \lambda \left\| \alpha_i^{t/b} \right\|_1 \quad (3)$$

where we set $\lambda = 0.15$ to control the sparsity. Note that $\alpha_i^{t/b}$ encodes superpixel x_i with the target/background dictionary. After that, we compute the SRE by

$$\varepsilon_i^{t/b} = \left\| x_i - D^{t/b}\alpha_i^{t/b} \right\|_2^2 \quad (4)$$

where ε_i^t and ε_i^b denote the T-SRE and B-SRE. Since we interpret high saliency as a combination of low T-SRE and high B-SRE, the superpixel-wise saliency map is defined as

$$Sal(s_i) = \rho^t \cdot \frac{1}{\varepsilon_i^t} + \rho^b \cdot \varepsilon_i^b \quad (5)$$

where we set weighting parameter $\rho^t = 1$ and $\rho^b = 2$. And the ε_i^t and ε_i^b are normalized to [0,1] in advance. Thus, we are able to estimate the location of airport ROI in the whole image according to the saliency map. After that, we employ the Otsu algorithm [15] that provides adaptive threshold to obtain a binary saliency mask M_s.

2.3. Target-aware Active Contour Model (TAACM)

In computer vision field, active contour model (ACM) has been extensively applied in medical and natural scene image segmentation. This approach describes the object contour as an evolving curve embedded in the artificially-designed energy functions, and pulls it towards the object edges by energy minimizing [16]. However, traditional ACMs are not fully practical in complex remote sensing scenes for being quite time-consuming in large-size RSIs. In addition, existing schemes can only separate visually prominent foregrounds from wide-range background, but often fail to further refine the target area from other redundant foreground objects. Illustrations of this problem are presented in Fig. 2.

Actually, we cannot distinguish the task-required target area with foreground redundancies with a totally data-driven formulation of the active contour model. Therefore, in this paper, we introduce saliency priors obtained by the SRS model into the original formulation of classic C-V model, as cited in [17], and put forward a novel target-aware active contour model (TAACM).

In the first stage, saliency priors obtained by SRS model are used to estimate the characteristics of airport area. For all pixel points $\Lambda = \{p_1, p_2, ..., p_m\}$ covered by the saliency mask M_s, we obtain the airport feature vector by

$$v_s = \frac{1}{m}\left(\sum_{p\in\Lambda} p^l, \sum_{p\in\Lambda} p^a, \sum_{p\in\Lambda} p^b, \sum_{p\in\Lambda} p^m \right) \quad (6)$$

where $p^{l/a/b/m}$ denote the *L*a*b* color values and gradient magnitude of a particular pixel. Obviously, this 4-d vector acts as an estimation of the airport features.

(a) (b) (c)

Fig. 2. *Unsuccessful segmentation using C-V [17] model where the airport is mixed up with foreground redundancies. (a) Input RSI, (b) and (c) Segmentation Result.*

For an image $U : \Omega \subset \Re^3 \to \Re$, and the corresponding luminance channel *I*, the TAACM consists of three items

$$E^{TAACM} = E^d + E^c + E^t \quad (7)$$

In formula (7), E^d denotes luminance differences between inside and outside the evolving contour. E^c controls curve smoothness. And E^t depicts the differences between areas estimated by saliency and areas incorporated in the contour curves. Specific expressions are given as

$$E^d = \int_\Omega (I - c_1)^2 H_\varepsilon(\psi) d\sigma + \int_\Omega (I - c_2)^2 \left[1 - H_\varepsilon(\psi)\right] d\sigma \quad (8)$$

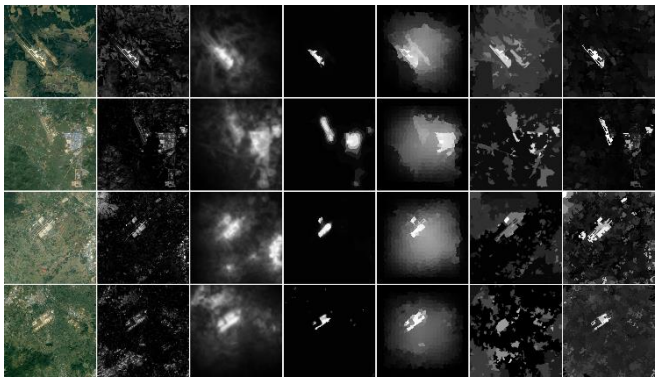
$$E^c = 0.3 \cdot \int_\Omega \delta_\varepsilon(\psi) \left|\nabla \psi\right| d\sigma \quad (9)$$

$$E^t = \int_\Omega \left\| v_c, v_s \right\|_2^2 \cdot H_\varepsilon(\psi) d\sigma \quad (10)$$

where ψ denotes level set function. c_1 and c_2 denote average intensities inside and outside the evolving contour. H_ε and δ_ε are regularized Heaviside and Dirac function. And v_c is the feature descriptor of the regions inside the contour curves. In practice, we minimize the level set energy functional with the Euler-Lagrange equation as

$$\frac{\partial \psi}{\partial t} = \delta_\varepsilon \left[0.3 div\left(\frac{\nabla \psi}{|\nabla \psi|}\right) - (I - c_1)^2 + (I - c_2)^2 - \left\| v_c, v_s \right\|_2^2 \right] \quad (11)$$

In the end, we present an acceleration scheme for a fast TAACM. On the one hand, we construct a local evolving window that centers the saliency mask where the energy evolution process is carried out. Since the adaptive window is typically much smaller, it will ignore most redundant information and largely enhance the speed of TAACM. On



(a) (b) (c) (d) (e) (f) (g)

Fig. 3. *Visual comparison of saliency maps with (a) Input, (b) FT, (c) CA, (d) PISA, (e) MR, (f) RC, and (g) SRS.*

the other hand, considering that the evolving curve may move very slowly if the initialization is far from the airport, the saliency mask will serve as the initial object contour.

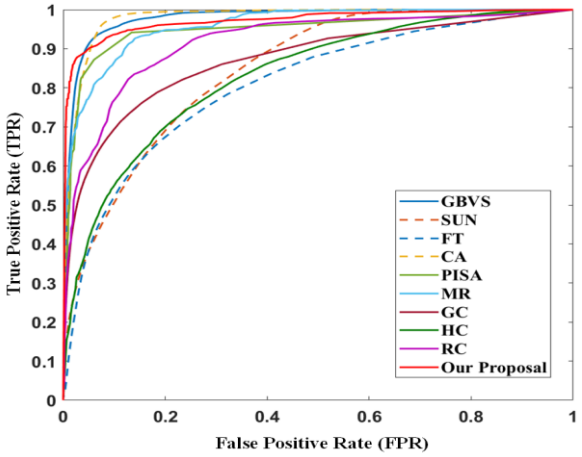


Fig. 4. *ROC curves of different saliency models.*

3. EXPERIMENT

We conducted an experiment to evaluate our proposal, using 230 images downloaded from Google Earth with 600×600 pixels. Another 42 samples are organized for the dictionary learning. The ground truth is manually labeled to finely depict structures instead of roughly covering boxes. All the competing models are run on MATLAB R2017b.

In the first stage, we compare the proposed SRS model with other nine state-of-the-art methods: GBVS [18], SUN [19], FT [20], CA [21], PISA [22], MR [23], GC [24], HC [25], and RC [25]. Due to space limitations, we only select five models for visual comparisons in Fig. 3. As it suggests, the SRS model not only highlights the airport ROI, but also largely suppresses the background. In Fig. 4, all competing saliency methods are evaluated by ROC curves, from which we can see that proposed SRS model fits the ground truth well for being close to the upper-left.

Besides, Fig. 5 illustrates the improvement of TAACM that significantly excludes redundant areas and refines the structures of airport target.

(a) (b) (c)

Fig. 5. *Successful segmentation using proposed TAACM where the airport is refined from foreground redundancies. (a) Input RSI, (b) and (c) Segmentation Result.*

In the end, we compare our proposal with work [2], [4], and [6] in terms of detection rate (DR), false alarm rate (FAR), and average running time. If the labeled area covers more than 50% of the ground truth, we consider it as a successful detection. Test data are given in Table 1.

Table 1. *Comparisons with Competing Methods*

Methods	DR	FAR	Time (s)
LF [2]	88.9%	18.9%	8.64
SM [4]	81.5%	25.2%	**5.22**
HRLS [6]	**92.4%**	14.6%	17.50
Ours	90.9%	**10.5%**	5.95

The LF [2] adopts an efficient feature representation based on Fisher vector to analyze line feature of runways and achieves relatively good efficiency. SM [4] uses an improved GBVS model for saliency detection and Hough transform for line detection. It is the fastest method, but is sensitive to non-airport line segments. HRLS [6] applies an iterative scheme to reinforce the image level by level, and achieves highest DR and low FAR. But the computational complexity largely increases due to the iteration process.

Comparatively, we achieve high DR and the lowest FAR. Since the dictionary is learned in advance and the TAACM is carried out within a local window, our proposal achieves very fast detection. Another major superiority of the ACM-based detection scheme lies in that it provides more detailed information of airport contours and structures.

4. CONCLUSION

An efficient airport extraction method is presented in this paper. First, we learn target and background dictionary in feature space. Second, a new sparse reconstruction saliency (SRS) model is proposed to detect salient regions in terms of sparse reconstruction error. Finally, we incorporate saliency priors and build up a novel target-aware active contour model (TAACM) that is more reliable in remote sensing scenes. Experiments indicate that the SRS model outperforms nine competing saliency algorithms in remote sensing scenes. And the airport extraction framework is faster and achieves higher detection accuracy and efficiency compared with other three detection methods.

5. REFERENCES

[1] Z. Kou, Z. Shi, and L. Liu, "Airport detection based on line segment detector," in *Conf. Computer Vision in Remote Sensing*. IEEE, 2012, pp. 72-77.

[2] B. Ümit, U. Halıcı, A. Şengür, et al., "Efficient Airport Detection Using Line Segment Detector and Fisher Vector Representation," *IEEE Geoscience and Remote Sensing Letters*, vol.13, no.8, pp. 1079-1083, 2016.

[3] Ö. Aytekin, U. Zöngür, and U. Halici, "Texture-based airport runway detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 3, pp. 471-475, 2013.

[4] X. Wang, Q. Lv, B. Wang, and L. Zhang, "Airport detection in remote sensing images: A method based on saliency map," *Cognitive neurodynamics*, vol. 7, no. 2, pp. 143-154, 2013.

[5] D. Zhu, B. Wang, L. Zhang, "Airport Target Detection in Remote Sensing Images: A New Method Based on Two-Way Saliency," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 5, pp. 1096-1100, 2015.

[6] D. Zhao, Y. Ma, Z. Jiang, et al., "Multiresolution Airport Detection via Hierarchical Reinforcement Learning Saliency Mode," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 6, pp. 2855-2866, 2017.

[7] L. Zhang, Y. Zhang, "Airport Detection and Aircraft Recognition Based on Two-Layer Saliency Model in High Spatial Resolution Remote-Sensing Images," *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, vol. 10, no. 4, pp. 1511-1524, 2017.

[8] L. Zhang, J. Zhang, "A new saliency-driven fusion method based on complex wavelet transform for remote sensing images," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 12, pp. 2433-2437, 2017.

[9] Q. Zhang, L. Zhang, W. Shi, Y. Liu, "Airport extraction via complementary saliency analysis and saliency-oriented active contour model," *IEEE Geoscience and Remote Sensing Letters*, 2018, doi: 10.1109/LGRS.2018.2828502.

[10] P. Zhang, X. Niu, Y. Dou, et al., "Airport Detection on Optical Satellite Images Using Deep Convolutional Neural Networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 8, pp. 1183-1187, 2017.

[11] T. Zhu, Y. Li, Q. Ye, et al., "Integrating saliency and ResNet for airport detection in large-size remote sensing images," in *Conf. Image, Vision and Computing*. IEEE, 2017, pp. 20-25.

[12] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274-2282, 2012.

[13] J. Mairal, F. Bach, J. Ponce, et al., "Online dictionary learning for sparse coding," *Proceedings of the 26th annual international conference on machine learning. ACM*, pp. 689-696, 2009.

[14] X. Li, H. Lu, L. Zhang, et al., "Saliency Detection via Dense and Sparse Reconstruction," *IEEE International Conference on Computer Vision*, pp. 2976-2983, 2013.

[15] L. Zhang, A. Li, X. Li, S. Xu and X. Yang, "Remote sensing image segmentation based on an improved 2-D gradient histogram and MMAD model," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 1, pp. 58-62, 2015.

[16] M. Gao, H. Chen, S. Zheng, et al., "A factorization based active contour model for texture segmentation," *IEEE International Conference on Image Processing*, pp. 4309-4313, 2016.

[17] T. Chan and L. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266-277, 2001.

[18] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," *in Proc. Adv. Neural Inf. Process. Syst.,* pp. 545-552, 2007.

[19] L. Zhang, M. H. Tong, T. K. Marks, H. Shan, and G. W. Cottrell, "SUN: A Bayesian framework for saliency using natural statistics," *Journal of Vision*, vol. 8, no. 7, pp. 32-32, 2008.

[20] R. Achanta, S. Hemami, F. Estrada, et al., "Frequency-tuned salient region detection," *IEEE Conference on Computer vision and pattern recognition*, pp. 1597-1604, 2009.

[21] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 10, pp. 1915–1926, 2012.

[22] K. Shi, K. Wang, J. Lu, et al., "PISA: Pixelwise Image Saliency by Aggregating Complementary Appearance Contrast Measures with Spatial Priors," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2115-2122, 2013.

[23] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang, "Saliency detection via graph-based manifold ranking," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3166–3173, 2013.

[24] M. M. Cheng, J. Warrell, W. Y. Lin, S. Zheng, V. Vineet, and N. Crook, "Efficient salient region detection with soft image abstraction," *IEEE Conference on Computer Vision*, pp. 1529–1536, 2013.

[25] M. M. Cheng, G. X. Zang, N. J. Mitra, X. Huang, and S. M. Hu, "Global contrast based salient region detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 569-582, 2015.