Airport Extraction via Complementary Saliency Analysis and Saliency-Oriented Active Contour Model

Qijian Zhang, Libao Zhang, *Member, IEEE*, Wenqi Shi, and Yue Liu

***Abstract*— Automatic airport extraction in remote sensing images (RSIs) has been widely applied in military and civil applications. An efficient airport extraction framework for RSIs is constructed in this letter. In the first step, we put forward a two-way complementary saliency analysis (CSA) scheme that combines vision-oriented saliency and knowledge-oriented saliency for the airport position estimation. In the second step, we construct a saliency-oriented active contour model (SOACM) for airport contour tracking, where a saliency orientation term is incorporated into the level-set-based energy functions. Under the guidance of saliency feature representations obtained by CSA, the SOACM can acquire well-defined and highly precise object contours. Experimental results demonstrate that the proposed extraction framework shows good adaptability in remote sensing scenes, and uniformly achieves high detection rate and low false alarm rate. Compared with three state-of-the-art algorithms, our proposal can not only estimate the location of airport targets, but also extract detailed information of the airport contours.**

***Index Terms*— Active contour model (ACM), airport extraction, object detection, remote sensing, saliency analysis.**

I. Introduction

With the remote sensing technology highly developed, automatically recognizing the airport targets in remote sensing images (RSIs) has become one of the most important but challenging computer vision problems. In reality, it is applied to airport navigation, military reconnaissance, and many other practical applications.

To the best of our knowledge, previous studies on this issue can be divided into two categories: unsupervised works that are based on airport feature modeling and works that introduce the supervised learning mechanism into the detection problem. The first scheme basically puts emphasis on the artificially designed representations of the airport geometrical characteristics, and the regions of interest (ROIs) can be recognized in terms of line segment detection [1], or saliency features that combine texture information [2], [3]. Since the first class uses primary properties of the airports, it can achieve fast detection with relatively good recognition results. However, as the best discriminative feature representations for the airport targets are difficult to predict and artificially construct, this method can be sensitive to complex background noises and the existence of irrelevant linear objects. On the contrary, in the second class, the supervision mechanism and machine learning are incorporated into the detection framework. In most cases, this method identifies the targets by means of an appropriate feature classifier, such as support vector machine (SVM) [4], and Adaboost algorithm [5]. In [6], the airport is described by a set of scale invariant feature transform keypoints and then selected from candidate regions by SVM.

In recent years, a lot of studies have been conducted from the perspective of image saliency analysis [7], where a variety of saliency cues are employed to take the place of traditional airport feature descriptions. In [8], the top-down and bottom-up saliency maps are combined to separate the candidate regions. Then, the airport ROI is determined with a pretrained SVM. Besides, the deep learning theory also receives more and more attention in image processing applications. For example, in [9] and [10], the convolutional neural network (CNN) is utilized to extract high-level features and hierarchical representations of the objects. With a well-designed learning network and training set, the CNN-based method uniformly presents highly robust detection results. In general, supervised detection algorithms achieve better recognition rate than do unsupervised methods. But it needs a large quantity of image samples that are precisely marked by the researchers, which makes it task-dependent and causes low model reusability. The process of pattern matching and sample training can be quite time consuming, and largely determines the quality of the whole detection framework.

In an image, the salient areas will catch human being's visual attention with low-level properties such as luminance, color, and contrast distribution. This vision-oriented saliency (VOS) indiscriminately extracts the eye-catching areas. For another, an observer can distinguish the airport targets by inherent prior cognition. This knowledge-oriented saliency (KOS) selectively focuses on the regions that contain more line segments, but fails to separate the airport runways from other disturbing linear objects such as the residential areas, long rivers, and highway.

Inspired by the two types of visual saliency mechanisms, this letter presents a complementary saliency analysis (CSA) model. The VOS connects contrast distribution with spatial relation of image subregions. The KOS considers the airport target as a cluster of organized lines and generates a quality-weighted line density map. To gain the information of airport contours, we incorporate fusion saliency representations into the framework of traditional active contour model (ACM), and formulate an innovative saliency-oriented ACM (SOACM).

Manuscript received October 29, 2017; revised February 20, 2018; accepted March 14, 2018. Date of publication May 2, 2018; date of current version June 22, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 41771407, Grant 61571050, and Grant 61071103, in part by Beijing Natural Science Foundation under Grant 4162033, and in part by the Open Fund of the State Key Laboratory of Remote Sensing Science under Grant OFSLRSS201621. *(Corresponding author: Libao Zhang.)*

The authors are with the College of Information Science and Technology, Beijing Normal University, Beijing 100875, China (e-mail: libaozhang@bnu.edu.cn).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2018.2828502

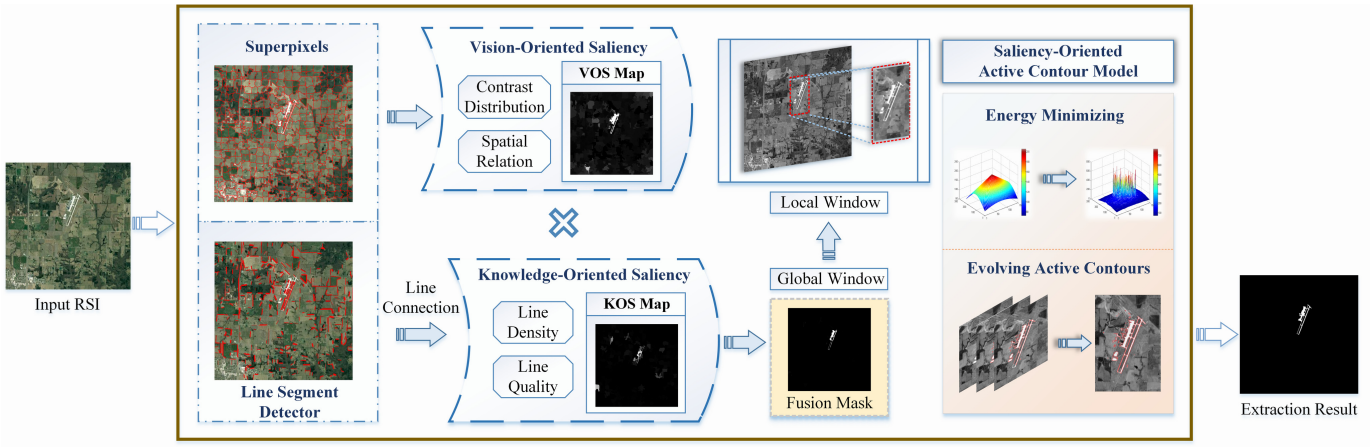

Fig. 1. Flowchart of the proposed airport extraction method.

In our implementation, a saliency orientation term (SOT) is added to the level-set-based energy functions, so that the object contour keeps evolving in the right direction and converges faster to the airport runways.

In conclusion, the contributions of this letter mainly lie in the following two aspects.

1) We put forward a CSA scheme that can accurately estimate the spatial location of the airport in remote sensing scenes.
2) We present an innovative SOACM for airport contour tracking. Compared with existing ACMs, the proposed SOACM significantly improves the evolving speed and succeeds in extracting airport contours that highly conform to the actual objects.

II. Model Construction

The proposed detection framework starts by segmenting the input RSI into a series of superpixels. In the first step, the VOS generates a bottom-up saliency map in terms of the interactions of interior contrast and spatial location among the superpixels. In the second step, the KOS detects the line segments and draws a top-down saliency map in terms of line density distribution. Finally, the proposed SOACM exploits the saliency feature representations to guide the process of energy minimizing, and acquires the airport contours within a small localized operating window. The flowchart of our method is shown in Fig. 1.

A. Superpixel Segmentation

Superpixel segmentation usually acts as a preprocessing step to reduce the complexity of subsequent image processing tasks. It decomposes the image into homogeneous subregions with well-preserved boundaries, aiming to simplify the image details and highlight structural information. In this letter, an efficient algorithm, simple linear iterative clustering (SLIC) [11], is introduced to acquire superpixels from RSIs. The SLIC method can generate regular and compact superpixels, with adjustable number of clusters. Compared with pixel-wise saliency models, the superpixel-based scheme can present more concise image information and is more robust to remote sensing noises.

Mathematical morphological filter is a useful tool for image noise suppression. Since superpixel segmentation is typically used for natural image processing, the quality of segmentation can be degraded in complex RSIs with large quantity of colors and uneven luminance distribution. To solve this, we utilize a morphology closing operator, where the image area will be first dilated and then be eroded, to strengthen the line segments of airport runways and remove small, isolated image fragments.

B. Complementary Saliency Analysis (CSA)

Human vision systems extract salient objects from complex background without effort. However, automatically detecting the salient region from images is still a challenging problem. In this section, we construct a two-way CSA model that is operated on superpixels. In the VOS layer, low-level saliency cues of contrast distribution and spatial relation are taken into account to acquire visually prominent candidate regions. In the KOS layer, we describe the airport target in terms of regional line density and length quality. The VOS and KOS are designed to be complementary, inter-reinforced, and highlight the airport ROI from different and independent aspects.

Note that in this letter, saliency is used particularly to guide the curve evolution of ACM and enhance the quality of airport contour tracking, not to directly extract the airport targets. Therefore, it would be acceptable if the CSA fails to acquire the complete ROI, as long as the fusion saliency mask does fall on part of the actual airport fields. And this enables us to impose stricter target matching criteria. Generally, the fusion saliency is given by

$$\mathrm{Sal}(r_k) = \mathrm{VOS}(r_k) \cdot \mathrm{KOS}(r_k) \tag{1}$$

where $\mathrm{VOS}(r_k)$ and $\mathrm{KOS}(r_k)$ mean the vision and knowledge layer saliency of the superpixel r_k.

C. VOS With Low-Level Saliency Cues

The VOS layer implements the assumption that the salient region distinguishes itself by higher contrast with other parts of the image. We also notice that the distance between two regions can largely determine to what extent one makes contrast with another. Thus, we naturally obtain the following statements.

1) Higher contrast is indicative of higher saliency.
2) Contributions that contrast make to saliency declines as the distance between two regions increases.

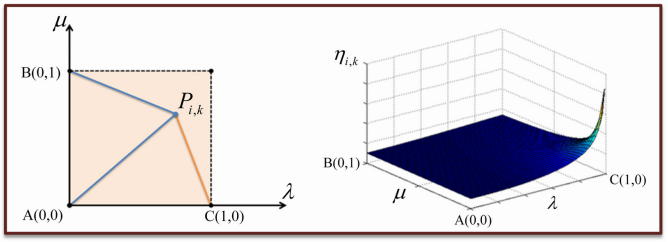

Fig. 2. Diagram of formulating the connection term in a coordinate system and its 3-D visual form.

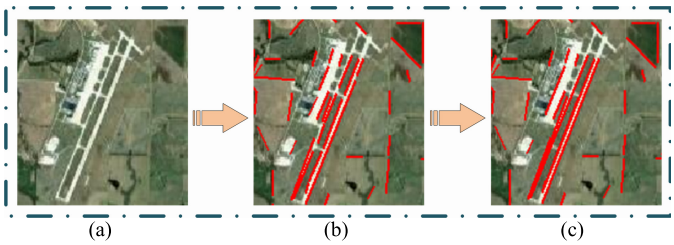

Fig. 3. Line segment connection. (a) Original image. (b) LSD results. (c) Connection results.

Based on the knowledge above, the contrast-based saliency for superpixel r_k can be formulated as follows:

$$\mathrm{CT}(r_k) = \sum_{r_i \neq r_k} D_c(r_i, r_k) \cdot \eta_{i,k} \tag{2}$$

where $D_c(r_i, r_k) = \|c(r_i), c(r_k)\|$ denotes mean color distances in $L \times a \times b$ color space between superpixel r_i and r_k. $\eta_{i,k}$ is a connection term that reflects the influence of area distances.

Specifically, let $\lambda = D_c(r_i, r_k)$ be the color differences and $\mu = \|\mathrm{center}(r_i), \mathrm{center}(r_k)\|$ be the region distance between r_i and r_k. Suppose that we apply the λ and μ in a coordinate system, with each normalized to [0, 1]. As shown in Fig. 2, any two regions can be mapped to a point $P_{i,k}$ within the colored area. The connection term is expected to reach its maximum at point C and decreases while approaching point A and B. To achieve this, we formulate $\eta_{i,k} = \eta_{i,k}(\lambda, \mu)$ as follows:

$$\eta_{i,k} = \sqrt{L_1 / L_2} \tag{3}$$

where L_1 and L_2 represent the geometrical distances as

$$L_1 = |P_{i,k}A| + |P_{i,k}B| = \sqrt{\lambda^2 + \mu^2} + \sqrt{\lambda^2 + (\mu - 1)^2} \tag{4}$$

$$L_2 = |P_{i,k}C| = \sqrt{(\lambda - 1)^2 + \mu^2}. \tag{5}$$

Under ideal circumstances where the object is segmented as a whole, we can get the ROI labeled equally salient. However, in most cases, the airport area is cut into pieces and consists of a cluster of adjacent superpixels. These superpixels may inhibit each other, causing uneven saliency distribution. Therefore, it is necessary to introduce a compensation operator so that the airport area can be uniformly highlighted.

Let $V = \{v_1, v_2, \ldots, v_m\}$ be a set of neighboring superpixels of r_k. We acquire the VOS with a local smoothing scheme as

$$\mathrm{VOS}(r_k) = \frac{1}{m+1} \cdot \left[\mathrm{CT}(r_k) + \sum_{i=1}^{m} \mathrm{CT}(v_i) \right]. \tag{6}$$

After that, we can acquire a series of candidate ROIs, from which the target area can be further distinguished by high-level prior knowledge of the airport objects.

D. KOS With High-Level Geometrical Priors

The KOS layer puts emphasis on geometrical characteristics of airport runways. It is observed that the airport area consists of a series of densely aggregated line segments. Although there inevitably exist many other disturbing linear objects, the airport area will distinguish itself with significantly larger line density.

Thus, the KOS of each superpixel can be measured on the basis of the following understanding.

1) A superpixel should be labeled more salient if it contains more line elements.
2) The line segments with higher quality should be given greater weight. In this letter, we naturally interpret the line quality as its length.

Line segment detector (LSD) [12] is an efficient linear-time model that outperforms existing line detection algorithms. Inspired by [7], the KOS layer first employs the LSD to extract line segments and generate the line density map. For superpixel r_k, we define the length-weighted line density as

$$\mathrm{LD}(r_k) = \frac{N_L(r_k)}{N(r_k)} \cdot Q_L(r_k) \tag{7}$$

where $N(r_k)$ is the number of pixels contained in region r_k and $N_L(r_k)$ counts the pixels on the line segments detected by LSD. Since the line segments that make up the airport runways are supposed to be long enough, we introduce an emphasis factor to represent the line quality. Obviously, superpixels that contain short and zigzag disturbing lines can be suppressed.

Let $L(k) = \{l_1, l_2, \ldots, l_p\}$ be a set of lines incorporated into r_k. Then, the emphasis factor $Q_L(r_k)$ is given by

$$Q_L(r_k) = \sum_{i=1}^{p} \mathrm{Length}(l_i). \tag{8}$$

Based on the length-weighted line density map, we formulate the KOS with a Gaussian form, so that the high-density areas can highly stand out ($\sigma^2 = 0.4$)

$$\mathrm{KOS}(r_k) = \exp\left\{ \frac{\mathrm{LD}(r_k)}{\sigma^2} \right\}. \tag{9}$$

In practice, the line segments of airport runways detected by LSD are always discontinuous, which makes it hard to separate them from other noise lines for being too short. To get better use of the line quality factor and distinguish the airport from other high-density regions, an effective line connection process is needed. Here, we employ a simple connection scheme stated in [8] to enhance the quality of line extraction. An example of this procedure is illustrated in Fig. 3.

E. Saliency-Oriented Active Contour Model (SOACM)

In computer vision field, ACM is extensively used for contour tracking and image segmentation. This approach describes the object contour as an evolving curve embedded in artificially designed energy functions, and pulls it toward the object edges by energy minimizing.

Compared with other detection frameworks, the conspicuous superiority of ours lies in that we combine saliency features with ACM and make further efforts in contour extraction. To illustrate the need for applying saliency orientation in

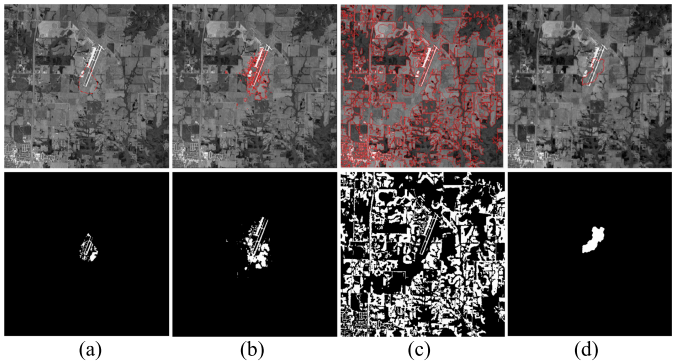

Fig. 4. Comparisons of the ACMs directly applied to the RSI. The first row visualizes the labeling curve. The second row shows the binary mask. Results obtained by (a) GVF [13], (b) C–V [14], (c) LBF [15], and (d) DRLSE [16].

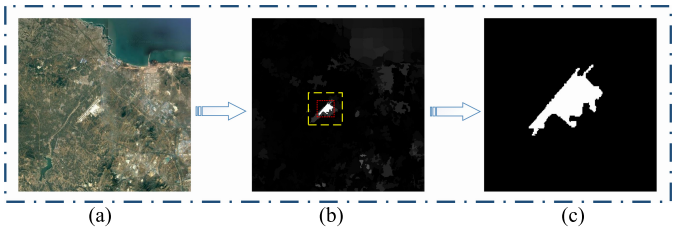

Fig. 5. Process of constructing the LEW and saliency-guided initialization of the proposed SOACM. (a) Original image. (b) Local evolving window. (c) Initial object contour.

ACM, an experiment is conducted using four popular ACMs. As shown in Fig. 4, in remote sensing scenes, the quality of contour recognition can be greatly degraded if the models are directly applied. Some of the algorithms become time consuming when the image size significantly increases.

Inspired by the C–V model [14], this letter innovatively puts forward an SOACM with better adaptability to RSIs. First, we construct a smaller local evolving window (LEW) that centers the saliency mask, in which the energy evolution is carried out. This helps to ignore redundant image context and guarantees high running speed. Considering that the evolving curve may move very slowly if the initialization is far from the airport area, we have the fusion saliency mask serve as the initial object contour. This avoids false alarm to a large extent. The process of constructing the LEW and model initialization is described in Fig. 5.

Mathematically, for a given image $U : \Omega \subset \Re^3 \to \Re$, and the corresponding 2-D gray scale image I, the level-set-based SOACM is formulated as

$$E^{\mathrm{SOACM}}(c_1, c_2, \psi) = \lambda_1 \int_\Omega (I - c_1)^2 H_\varepsilon(\psi) d\sigma + \lambda_2 \int_\Omega (I - c_2)^2 [1 - H_\varepsilon(\psi)] d\sigma + \alpha \int_\Omega \delta_\varepsilon(\psi) |\nabla \psi| d\sigma + E_{\mathrm{ort}}(c_1) \quad (10)$$

where ψ means the level set function. c_1 and c_2 denote average intensities inside and outside target contour. H_ε and δ_ε are regularized Heaviside and Dirac function. λ_1 and λ_2 are set to 1 and α is set to 0.2.

However, to identify the target from the whole image context with two average values can be unreliable in complex RSIs. In most cases, there should exist more than one visually prominent foregrounds that C–V fails to make further selections. Thus, the airport will be mixed up with foreground redundancies. In order to refine and pinpoint airport ROI, we supplement an SOT E_{ort} in the energy formulation as

$$E_{\mathrm{ort}}(c_1) = \lambda_3 \int_\Omega (U - CV_{\mathrm{ort}})^2 H_\varepsilon(\psi) d\sigma \quad (11)$$

where λ_3 is set to 1. CV_{ort} is a predetermined prior feature vector, referring to the mean color value of all pixels covered by the obtained saliency mask. The SOT would decrease if more pixels with similar colors to the airport ROI are incorporated, and increase as the curve moves toward those with different colors, not just the background elements. In practice, we minimize the level-set energy functional with the Euler–Lagrange equation as

$$\frac{\partial \psi}{\partial t} = \delta_\varepsilon \left[\alpha \mathrm{div}\left(\frac{\nabla \psi}{|\nabla \psi|} \right) - \lambda_1 (I - c_1)^2 + \lambda_2 (I - c_2)^2 - \lambda_3 (U - CV_{\mathrm{ort}})^2 \right]. \quad (12)$$

On the basis of reliable saliency priors obtained by CSA, the evolving curve will be guided to approach the target. In the end, we remove all small fragments of image islets.

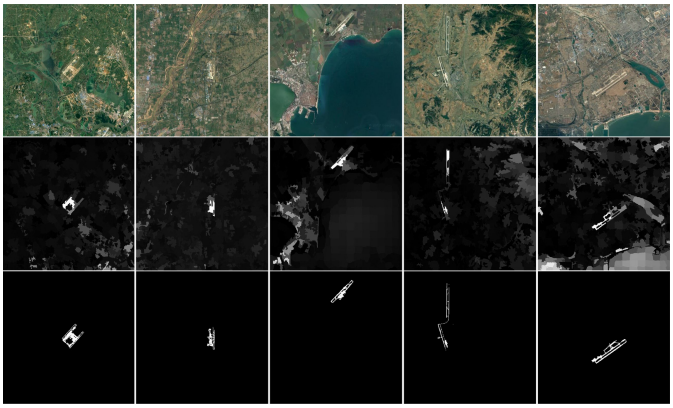

Fig. 6. Saliency maps and results of airport extraction obtained by the proposed CSA and SOACM.

III. Experimental Results

An experiment is performed to evaluate our framework. We utilize 276 RSIs gathered from Google earth with the resolution ranging from 20 to 40 m. The size of images is 600 × 600 pixels. The ground truth is manually labeled to finely depict the airport structures and details.

Fig. 6 presents saliency maps and detection results obtained by CSA and SOACM. As it shows, the saliency model succeeds in providing a rough but reliable estimation of the airport's space location. The SOACM adapts well to RSIs and generates complete, detailed, and highly accurate airport contours. The proposed framework is evaluated by detection rate (DR), false alarm rate (FAR), mean absolute error (MAE), and running time. In practice, if the extracted area contains more than 50% of the ground truth, we consider it as a successful detection. Besides, MAE is defined to measure average pixel differences between the detection results and

TABLE I
PERFORMANCES OF DIFFERENT AIRPORT DETECTION MODELS

Methods	DR	FAR	MAE	Time (s)
Ref [2]	80.2%	26.9%	0.073	**5.32**
Ref [4]	88.5%	15.7%	0.051	8.27
Ref [7]	**93.6%**	**12.9%**	0.015	17.30
Ours	91.7%	14.5%	**0.007**	9.75

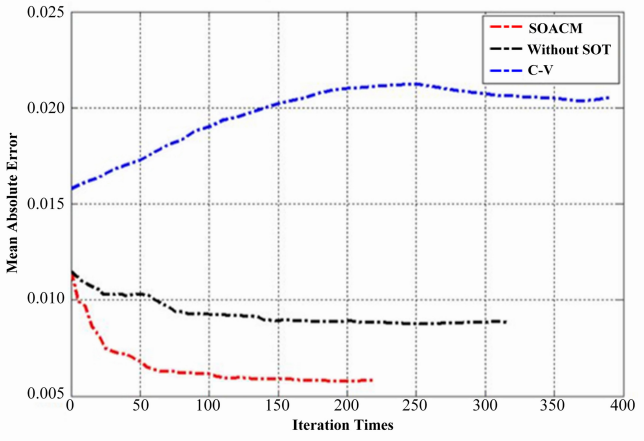

Fig. 7. MAE of the extracted airport contours using different methods.

ground truth as

$$\mathrm{MAE} = \frac{1}{W \times H} \sum_{x=1}^{W} \sum_{y=1}^{H} |R(x, y) - GT(x, y)| \qquad (13)$$

where W and H denote the image width and height, $R(x, y)$ is the pixel value, and $GT(x, y)$ is the ground truth.

We compare our model with three different types of methods: one that uses saliency analysis [2], a line detection-based method that combines Fisher vector coding and SVM [4], and one that applies the idea of machine learning [7]. The test data are given in Table I.

As it shows, our proposal achieves higher DR and lower FAR than do works [2] and [4]. Compared with [7], we achieve comparable accuracy and reliability, and our algorithm is much faster and omits the time-consuming work of sample marking and model training. Moreover, although the DR is slightly lower than that in [7], our proposal shows much better MAE, which means that we are able to provide a more precise and detailed description of the airport contours.

In Fig. 7, performances on MAE that change as the iteration times increase are investigated using the C–V and SOACM. Moreover, a scheme without the SOT is tested. We can see that the original C–V shows increasing MAE for it incorporates too much noise elements into the evolving contour. Instead, the curve of the scheme without SOT keeps decreasing, starts at a lower level of MAE, and moves faster. Because the initial object contour is determined by the saliency mask, and the energy evolution is operated within the LEW, which is much smaller compared with the input RSI. By contrast, the SOACM achieves the lowest MAE with the minimum number of iterations, for the reason that the evolving curve is guided by the SOT to approach the airport area.

IV. CONCLUSION

This letter focuses on airport extraction in RSIs. A two-way CSA model is proposed to estimate airport ROI by combining low-level VOS and high-level KOS cues. Based on the fusion saliency map, we present a novel SOACM to acquire the airport contours. Considering that conventional ACMs are less reliable in RSIs with faint contrast and ill-defined boundaries, we construct an SOT to guide the evolution of energy functions with saliency feature representations. We compare our model with three state-of-the-art methods in terms of DR, FAR, and MAE measure. Experimental results prove that our model can not only achieve high detection accuracy and efficiency, but also extract contours that highly conform to the actual objects.

REFERENCES

[1] Z. Kou, Z. Shi, and L. Liu, "Airport detection based on line segment detector," in *Proc. Int. Conf. Comput. Vis. Remote Sens.*, Dec. 2012, pp. 72–77.
[2] X. Wang, Q. Lv, B. Wang, and L. Zhang, "Airport detection in remote sensing images: A method based on saliency map," *Cognit. Neurodyn.*, vol. 7, no. 2, pp. 143–154, 2013.
[3] L. Zhang and Y. Zhang, "Airport detection and aircraft recognition based on two-layer saliency model in high spatial resolution remote-sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 4, pp. 1511–1524, Apr. 2017.
[4] U. Budak, U. Halıcı, A. Şengür, M. Karabatak, and Y. Xiao, "Efficient airport detection using line segment detector and Fisher vector representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 8, pp. 1079–1083, Aug. 2016.
[5] O. Aytekin, U. Zongur, and U. Halici, "Texture-based airport runway detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 471–475, May 2013.
[6] C. Tao, Y. Tan, H. Cai, and J. Tian, "Airport detection from large IKONOS images using clustered SIFT keypoints and region information," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 1, pp. 128–132, Jan. 2011.
[7] D. Zhao, Y. Ma, Z. Jiang, and Z. Shi, "Multiresolution airport detection via hierarchical reinforcement learning saliency model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 6, pp. 2855–2866, Jun. 2017.
[8] D. Zhu, B. Wang, and L. Zhang, "Airport target detection in remote sensing images: A new method based on two-way saliency," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 5, pp. 1096–1100, May 2015.
[9] Z. Xiao, Y. Gong, Y. Long, D. Li, X. Wang, and H. Liu, "Airport detection based on a multiscale fusion feature for optical remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 9, pp. 1469–1473, Sep. 2017.
[10] F. Zhang, B. Du, L. Zhang, and M. Xu, "Weakly supervised learning based on coupled convolutional neural networks for aircraft detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 9, pp. 5553–5563, Sep. 2016.
[11] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.
[12] R. G. von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, "LSD: A fast line segment detector with a false detection control," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 4, pp. 722–732, Apr. 2010.
[13] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 359–369, Mar. 1998.
[14] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.
[15] C. Li, C.-Y. Kao, J. C. Gore, and Z. Ding, "Implicit active contours driven by local binary fitting energy," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–7.
[16] C. Li, C. Xu, C. Gui, and M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3243–3254, Dec. 2010.