An Edge-Boxes-Based Intruder Detection Algorithm for UAV Sense and Avoid System

ZHANG Zhouyu[1], *CAO Yunfeng*[1*], *ZHONG Peiyi*[1], *DING Meng*[2], *HU Yunqiang*[3]

1. College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, P.R.China;

2. College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, P.R.China;

3. College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, P.R.China

Abstract: With the great development of unmanned aircraft system (UAS) over the last decade, sense and avoid (SAA) system has been a crucial technology for integrating unmanned aircraft vehicle (UAV) into national airspace with reliable and safe operations. This paper mainly focuses on intruder detection for SAA system. A robust algorithm based on the combination of edge-boxes and spatial pyramid matching using sparse coding (sc-SPM) is presented. The algorithm is composed of three stages. First, edge-boxes method is adopted to obtain a large number of proposals; Second, the optimization program is presented to obtain intruder area-of-interest (ROI) regions; Third, sc-SPM is employed for feature representation of ROI regions and support vector machines (SVM) is adopted to detect the intruder. The algorithm is evaluated under different weather conditions. The recall reaches to 0.95 in dawn and sunny weather and 0.9 in cloudy weather. The experimental results indicate that the intruder detection algorithm is effective and robust with various weather under complex background.

Key words: detection; unmanned aircraft vehicle (UAV); sense and avoid (SAA); edge-boxes; sc-SPM; ROI

CLC number: TP37   **Document code:** A   **Article ID:** 1005-1120(2019)02-0253-11

0 Introduction

Unmanned aerial vehicle (UAV) has shown great capacity in both military and civil aspects, so various UAVs are researched and manufactured by different countries. There is potential safety hazard for national airspace with the increasing number of UAVs[1]. With the widespread use of UAVs, how to integrate UAVs safely into national airspace has become an urgent problem. The sense and avoid technologies of UAVs are the most challenging and critical technique to solve this problem. SAA system has been identified as one of crucial technologies that UVAs can integrate into national airspace, and vision-based SAA system is regarded as a potential solution for maintaining the safety of national airspace[2].

The vision-based sense and avoid system includes two parts. Part 1 is sensing: Detecting and tracking potential collision aircrafts; Part 2 is avoiding: Avoiding maneuver to resolve the collision. The most important part of detection is analyzing the image information from visual sensors to locate the intruder. There is no doubt that detection is very important as the first stage of SAA system, because the result of detection decides directly whether the follow works can continue or not. So the stage of detection plays a quite important role in SAA system[2-4].

1 Related Works

In this paper, the images of visual sensor are used to study the problem of intruder detection. Difficulties and challenges of intruder detection are as follows: (1) Intruder detection with the complex background; (2) intruder detection with different

lighting conditions. An edge‑boxes‑based intruder detection algorithm is proposed for UAV SAA system, which is illustrated in Fig. 1 below. At the first, we utilize edge map to get massive proposals from original images, and obtain the appropriate regions‑of‑interest (ROIs) by optimizing proposals. Next, we extract feature from the ROI areas based on linear spatial pyramid matching by using sparse coding with the over‑complete dictionary. At last,

support vector machine (SVM) is trained based on sparse coding, and then used to identify the intruder from ROIs.

Fig.1 demonstrates that ROIs is the key part of the proposed algorithm. There is a traditional detection method that is based on sliding window[5-7]. However, sliding window method has a very poor real‑time performance. Therefore, the concept of ROI has been proposed[8-9].
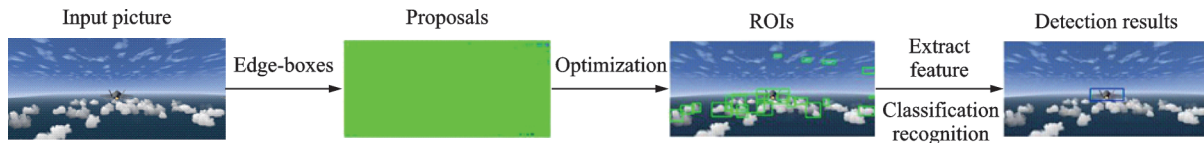

Fig.1   Block schema of the proposed algorithm

There are many ROI methods based on significance analysis, and the most typical methods include IT[10], MZ[11], LC[12], GB[13], SR[14], AC[15], FT[16], CA[17], HC[18] and RC[18]. The principle of IT and GB are based on biological structure, leading to a complex and time‑consuming calculation. The principle of other methods are based on calculating contract of color, texture, light and so on. Significant performance[19-23] is decided on differences in contrast characteristics, in other words, the smaller contract between the detection target and the background, the worse performance of those methods. In this paper, the images of potential intruder obtained by visual sensors often have low significance and obscured challenges because of complex backgrounds, bad weathers. Significant detection method cannot extract good ROIs, because the contrast between the target and background is too low. In order to solve this problem, edge‑boxes has been proposed by Piotr et al, using edge information to locate object proposals[24-25]. This algorithm can quickly and accurately locate the object proposals without learning, and is not affected by the size of the image, the complexity of the background.

After ROIs are obtained, the next work is to identify which ROI contains the detection target. The classical recognition algorithm consists of two parts：Feature extraction and classifier design. In terms of feature learning, human visual perception

system is worth learning. In order to simulate the visual perception mechanism of human brains, Hinton et al. proposed the concept of unsupervised feature learning in 2006[26-28]. Different from traditional shallow learning, unsupervised feature learning emphasizes the depth of the model structure and transforms the feature representation of original space into a new feature space by layer‑by‑layer feature transformation, making classification or prediction easier. This paper intends to use the unsupervised feature learning model based on sparse coding theory to deal with the complex background of intruder detection. The sparse coding method[29] has achieved success in many fields, like blind source signal separation, speech signal processing, natural image feature extraction, and pattern recognition. It is promising in our study.

2   Detailed Description of the Proposed Algorithm

Edge‑Boxes is an effective approach for increasing the computational efficiency of intruder detection. However, the number of bounding boxes will significantly increase when enough objects exist with complete profiles inside the image. For test images, there are a number of objects with full profile, so that the number of proposals is very huge, usually about 10 000. The proposed intruder detection algorithm is based on sc‑SPM which has good robust-

ness and accuracy but time-consuming. In order to improve the poor real-time performance of the slide window method, we need to dealing with massive proposals to select a small number of ROIs with appropriate size and small amount.

2. 1   Obtaining proposals based on edge-boxes

In this section, there are two stages to obtain ROI areas based on edge-boxes. The first stage is to generate object bounding box proposals using edge information. The second stage is to deal with the massive proposals to get appropriate ROIs.

First, edge image is obtained based on structured edge detection theory which effectively can predict object boundaries. The edge image has very good performance in edge presentation, but the edge is very dense. In order to get sparse edge image, we need to use non-maximal suppression (NMS) orthogonal to reduce the density of edge image.

Second, the edge points of a near straight line are concentrated in an edge group based on a greedy approach. Thus, we will have more than N edge groups. And then we need to compute the affinity between each pair of neighboring groups

$$\alpha(s_i, s_j) = \left| \cos(\theta_i - \theta_{ij}) \cos(\theta_j - \theta_{ij}) \right|^{\gamma} \quad (1)$$

where θ_{ij} is the angle between a pair of groups. If the pair of edge groups in a straight line, the similarity of the above formula is higher.

Third, we need to compute the value of every edge group

$$\omega_b(s_i) = 1 - \max_T \prod_j^{|T|-1} \alpha(t_j, t_{j+1}) \quad (2)$$

where T is an order path of edge groups that begins with some $t_1 \in S_b$ and end with $t_{|T|} = s_i$.

The edge group with the value of 1 is classified as part of the contour on the box, and the edge group with the value of 0 is classified as part of outside the box or the border overlap contour.

Forth, using the value of ω_b to define score

$$h_b = \frac{\sum_i \omega_b(s_i) m_i}{2(b_w + b_h)^k} \quad (3)$$

where b_w and b_h are the width and height of the box. If the score is high, the whole object contour is like-

ly to be included in the box.

2. 2   Optimizing proposals to obtain ROIs

We can attain tens of thousands of proposals by using edge-boxes method, which almost cover the whole picture densely, such as Fig.2(a). However, how to find the effective boxes among them is a crucial problem. Since the size of these massive proposals is not the same, there are some boxes that are big enough to cover almost the entire picture, and some boxes that are small enough to cover only a few pixels. These boxes are disadvantageous to detect the intruder, so we should build a set of optimization programs to remove ineffective proposals.

First, the score of each proposal can be obtained according to Eq.(3). Thus, the first step of optimization programs is to remove proposals with low scores. Different pictures have different distributions of scores, while share a similar small proportion of high scores. So an appropriate proportional threshold will work and reduce runtime. We boldly speculate that the appropriate threshold (called S) is between 0.01 and 0.1 based on experimental experiences. As a result, the performance of our algorithm with the threshold $S=0.03$ is ideal. An example is illustrated in Fig.2(b).

Second, an appropriate threshold should be set for retaining the appropriate size proposals. Too large ROIs would make the location accuracy of detection results low, so too big proposals should be eliminated. We boldly speculate that the appropriate threshold (called Z) is between 0.2 and 0.6 based on experimental experience. As a result, the performance of our algorithm with the threshold $Z=0.35$ is ideal. An example is illustrated in Fig.2(c).

Finally, there are appropriate proposals and too small proposals with high scores after the above two steps. These proposals are characterized by small border areas, but more intense in the complete profile. We perform non-maximal suppression (NMS) orthogonal to solve this problem. This process is to find the optimal proposals by merging the overlapping proposals and removing the bad proposals. Thus we can get the sparse boxes including the

complete contour of object. These optimal proposals are the ROI areas, as depicted in Fig.2(d).
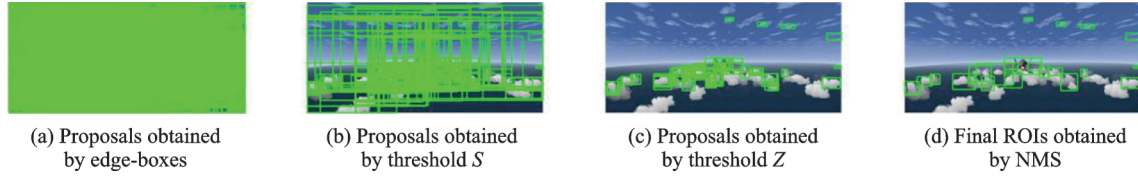
The whole optimization process is illustrated in Fig.2.

(a) Proposals obtained by edge-boxes   (b) Proposals obtained by threshold S (c) Proposals obtained by threshold Z (d) Final ROIs obtained by NMS

Fig.2   Optimization process

2.3 Extracting features and design classification

The feature extraction method based on linear spatial pyramid matching using sparse coding (sc-SPM) is composed of three stages.

(1) Extracting low-level features of ROI areas using SIFT.

(2) Sparse coding of low-level features based on the over-complete dictionary. For each image represented as a descriptor set Y, the SC codes are obtained by optimizing Eq. (4)

$$\min_{X} \ \| Y - DX \|_{F}^{2} \quad \text{s.t.} \quad \forall i, \|x_i\|_0 \leqslant T_0 \quad (4)$$

where D is the over-complete dictionary trained by K-SVD; T_0 the maximum value of non-zero entries of the coefficient vector. The optimization problem is convex in X (with D fixed). Coefficients (x) is computed with L1 norm regularization to be sparse, well known as Lasso in the Statistical literature.

(3) Max-pooling sparse vectors based on linear SPM. The sparse codes of the above process is represented as a matrix rather than vector, which cannot be adopted directly of the linear SVM classifier.

And then if the sparse codes are too large, the calculation will be more difficult and there will be an over-fitting problem. Therefore we need a pre-chosen pooling function to solve these problems. The max-pooling function is more robust than average pooling function for local spatial variations, so we defined the pooling function as a max pooling function

$$Z_j = \max \left\{ |x_{1j}|, |x_{2j}|, \cdots, |x_{Mj}| \right\} \quad j = 1, 2, \cdots, N \quad (5)$$

where x_{ij} is the element of the matrix $X \in \mathbf{R}^{M \times N}$, M is the number of N dimensional sparse vectors.

Since the features obtained by sc-SPM is non-linear, we can get a good performance using linear SVM classifier for identifying the intruder. For linearly separable data, discriminant function is set as

$$y_i(\boldsymbol{w}^{\mathrm{T}} \boldsymbol{x}_i + b) - 1 \geqslant 0 \quad i = 1, 2, \cdots, N \quad (6)$$

where $\boldsymbol{x}_i$ is an input linear sample, $y_i \in \{1, -1\}$ is the class label of $\boldsymbol{x}_i$.

3 Experiments and Results

3.1 Expermental setup

In the experiments, the test images are obtained by the simulative flight software Flight-Gear V3.4. For the diversity of the samples, we select massive images of F35 under three kinds of weather conditions with complex background.

The process of training linear SVM classifier include two steps.

Step 1 Extracting positive features of 400 images from 500 positive images and negative features of 600 images from 700 negative images. The rest images are test images.

Step 2 Obtaining classifier parameters by training the positive and the negative features using classifier function.

3.2 Results and analyses

In the experiment, there are two problem that we need to solve: How to judge the performance of the experimental results, and how to achieve optimal performance. For judging the experimental results, we need to introduce some indicators. For achieving optimal performance, we need to discuss some parameters to find the optimal results.

3.2.1 Performance indicators

We adopt recall as the performance indicators to evaluate the quality of the results. Recall rate is

the ratio of the number of detected targets and the number of the targets in the test library. Before we calculate the recall rate of detection results, we need to determine the value of an essential parameter: Intersection over union (IoU) which decides whether the detection results are correct or not. In the evaluation system of target detection, IoU is a crucial parameter that, in a simple term, is the overlap of the target window and the original mark window. The accuracy of a bounding box is typically measured using IoU metric. IoU computes the intersection of a detection box and the ground truth box divided by the area of their union.

3.2.2 Comparative experiments

The experimental results of optimized edge-boxes method have been compared with spectral residual (SR) method. SR method is the earliest and the most classic significance analysis method. Many significance analysis methods were proposed based on this method. Our method has compared with SR method from two ways. First, we compare the number of ROIs obtained by our method and SR method from the same image; Second, we compare the value of IoU from the ROIs of above methods. We randomly selected 100 pictures in every weather image library as test samples, and the experiments are illustrated in Fig.3.
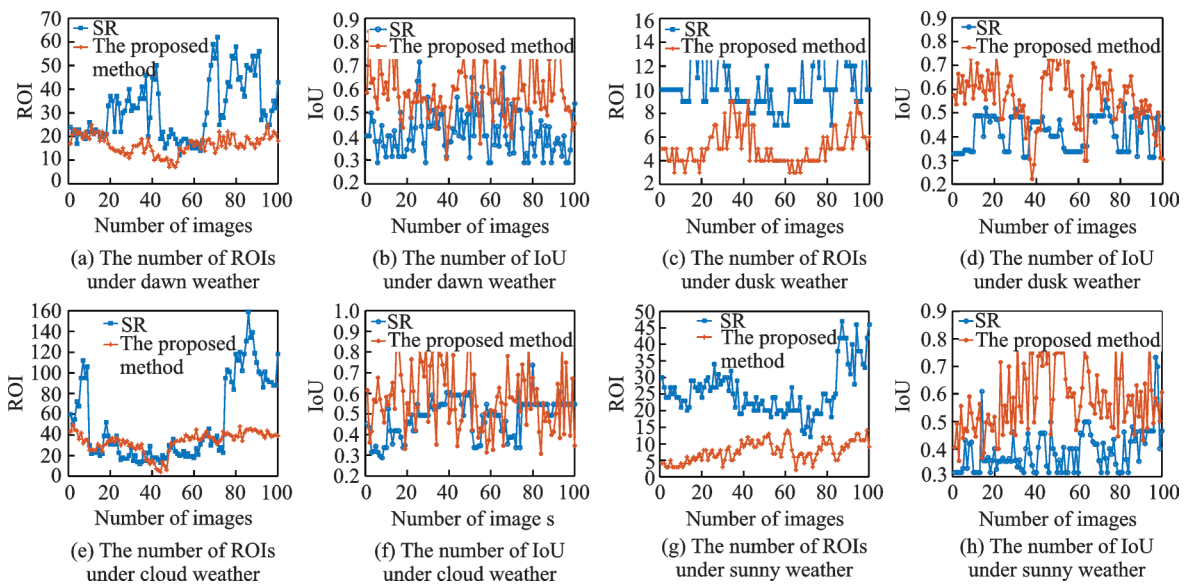

Fig.3   Comparative results under four weathers

Experiments show that the number of ROIs obtained by edge-boxes method is less than the one by SR method under different weathers, and the IoU of ROIs by edge-boxes method is more than the IoU of ROI by SR method under different weathers. The experimental performance of edge-boxes method is better than SR method.

3.2.3 Experimental results

There are four parameters that affect the final results, α, β, S, Z. α indicates the IoU for neighboring candidate bounding boxes to control the step size of the sliding window, and β controls the NMS threshold to merge sorted boxes for reducing density of proposals. S is the one maintaining the edge boxes with high scores, and Z can eliminate too big or too small edge boxes.

In order to reflect the diversity of the subjects, we choose four typical weathers with complicated background, including sunny, dawn, cloudy, dusk weathers. Then we randomly selected 100 pictures in every weather image library as test samples.

The experiments begin by testing our approach on various values of the four parameters $\alpha=0.65$, $S=0.03$, $Z=0.35$, $\beta=0.50$, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95. The recall which changes as the β changes can reflect whether the detected results are good or not. Fig.4 illustrates the influence of β to the intruder extraction algo-

rithm. As β is increased from 0.5 to 0.95, we can easily see that the value of recall changes, and reaches peaks in the process. These peaks indicate the best performance in different weather. When β is 0.75, the recall of dawn and sunny weathers will achieve the peak. However, the peak of the recall in the cloudy weather will be achieved when β is 0.6, and so will the peak of the dusk weather when β is 0.7.

According to Fig.4, our second set of experiments tests several variants of the algorithm. In these experiments, parameters should be set as: $\beta=$ 0.70, $S=0.03$, $Z=0.35$, $\alpha=0.45$, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90 in the dusk weather; $\beta=0.75$, $S=0.03$, $Z=0.35$, $\alpha=$ 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90 in the dawn and sunny weathers, and $\beta=$ 0.60, $S=0.03$, $Z=0.35$, $\alpha=0.45$, 0.50, 0.55,

0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90 in the cloudy weather. Fig.5 shows the recall varying with α from 0.45 to 0.90. As α is increased from 0.45 to 0.90, we can easily see that the value of recall changes and reaches peaks in the process. These peaks indicate the best performance in different weathers. When α is 0.65, the recall of dusk, dawn and sunny weathers will achieve peaks. However, the peak of the recall in the cloudy weather will be achieved when α is 0.7.

In Figs.4, 5, when $\alpha=0.65$, $\beta=0.75$, the recall rate is the highest in the dawn and sunny weathers, but the recall rate is the highest when $\alpha=0.70$, $\beta=0.60$ in the cloudy weather, and $\alpha=0.65$, $\beta=$ 0.70 in the dusk weather. Based on these data, we begin test our approach on the various values of the parameter S and Z.
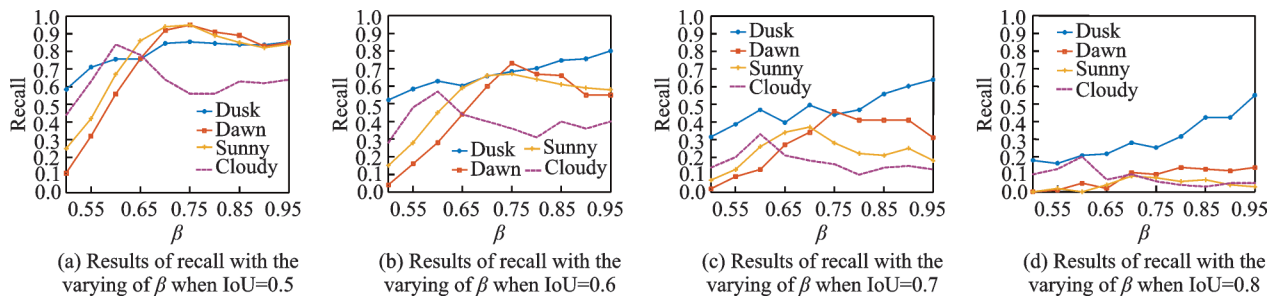

(a) Results of recall with the varying of β when IoU=0.5
(b) Results of recall with the varying of β when IoU=0.6
(c) Results of recall with the varying of β when IoU=0.7
(d) Results of recall with the varying of β when IoU=0.8

Fig.4   Comparison of recall changing with β by our algorithm in the four weathers

(a) Results of recall with the varying of α when IoU=0.5
(b) Results of recall with the varying of α when IoU=0.6
(c) Results of recall with the varying of α when IoU=0.7
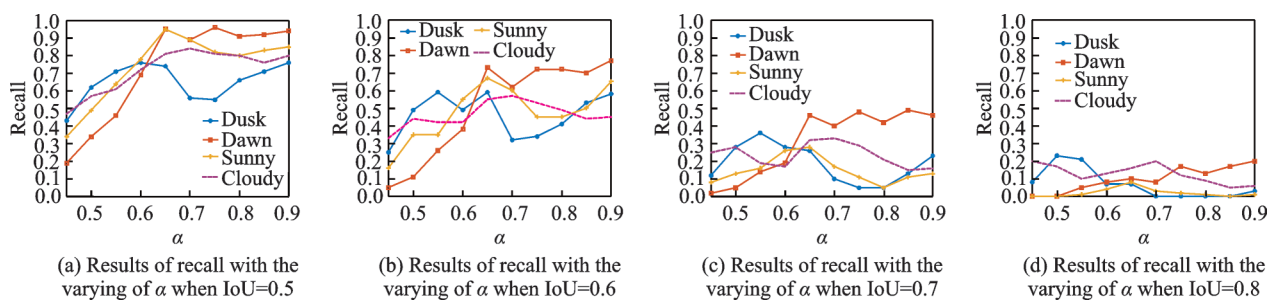(d) Results of recall with the varying of α when IoU=0.8

Fig.5   Comparison of recall changing with α by our algorithm in the four weathers

In the third experiment, parameters should be set as $\beta=0.70$, $\alpha=0.65$, $Z=0.35$, $S=0.01$, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, in the dusk weather; $\beta=0.75$, $\alpha=0.65$, $Z=0.35$, $S=0.01$, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, in the dawn and sunny weathers; and $\beta=0.60$, $\alpha=0.70$, $Z=0.35$, $S=0.01$, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, in the cloudy weather. The experimental results are il-

lustrated in Fig.6. As S is increased from 0.2 to 0.6, we can easily see that the value of recall changes and reaches peaks in the process. These peaks indicate the best performance in different weathers. When S is 0.03, the recall of dawn and sunny weathers will achieve peaks. However, the peak of the recall in the cloudy weather will be achieved when S is 0.04 and the peak of the recall in the dusk weather will be achieved when S is 0.02.
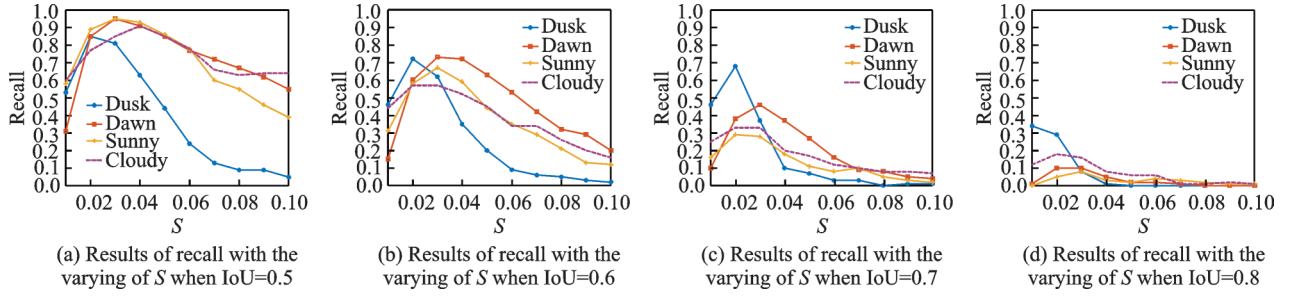
Fig.6   Comparison of recall changing with S by our algorithm in the four weathers

According to the third experiment, the forth experimental parameters are set as $\beta=0.70$, $\alpha=0.65$, $S=0.02$, $Z=0.20$, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60 in the dusk weathers; $\beta=0.75$, $\alpha=0.65$, $S=0.03$, $Z=0.20$, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60 in the dawn and sunny weathers; and $\beta=0.60$, $\alpha=0.70$, $S=0.04$, $Z=0.20$, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60 in the cloudy weather. The experimental results are illustrated in Fig.7. As Z is increased from 0.01 to 0.1, we can easily see that the recall changes and reaches peaks in the process. These peaks indicate the best performance in different weathers. When Z is 0.35, the recall of the four kinds of weathers will achieve peaks.
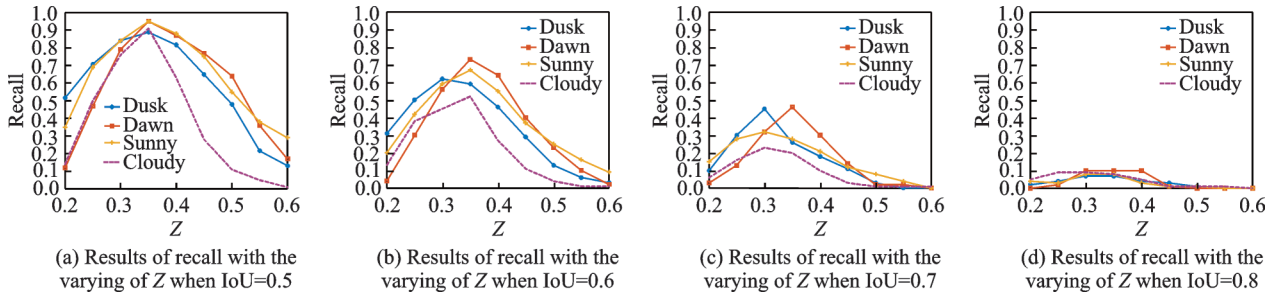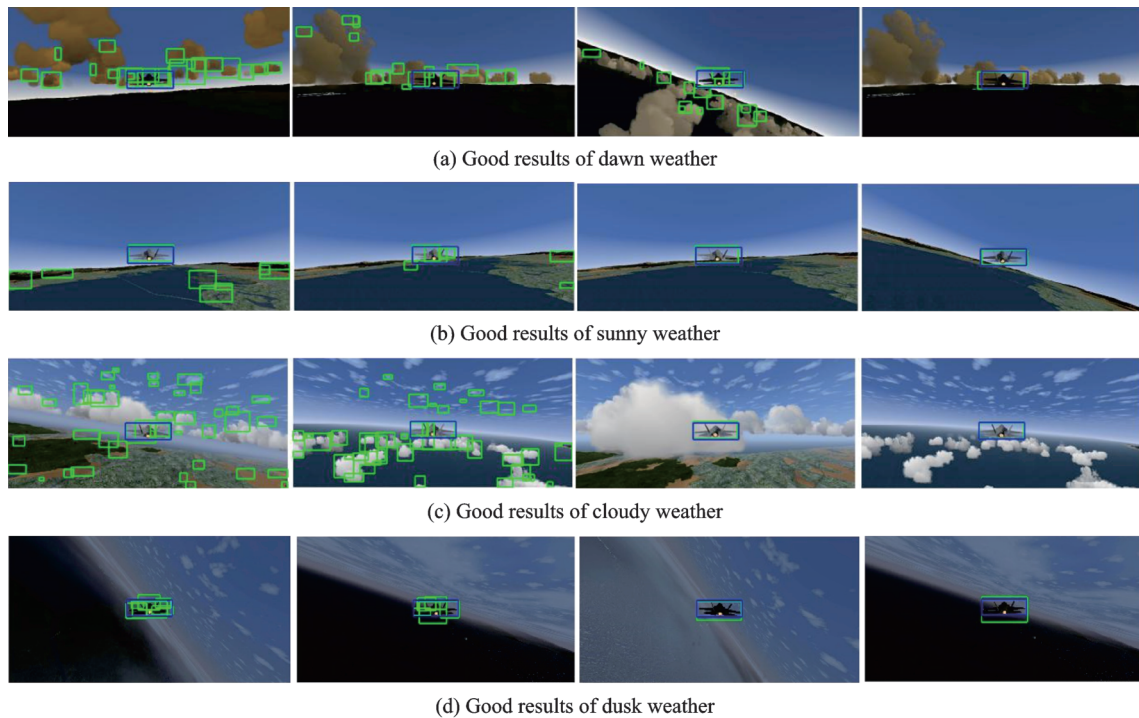

Fig.7   Comparison of recall changing with Z by our algorithm in the four weathers

From the above, when $\alpha=0.65$, $\beta=0.75$, $S=0.03$, $Z=0.35$ in the dawn and sunny weathers, the recall reaches peaks, which means the detection results are the best. But when in the cloudy weather, $\alpha=0.70$, $\beta=0.60$, $S=0.04$, $Z=0.35$, the recall reaches peaks, which means the detection results are the best; and when in the dusk weather, $\alpha=0.65$, $\beta=0.70$, $S=0.02$, $Z=0.35$, the detection results are the best. Some good results of different weathers are illustrated in Fig.8, and the bad results are illustrated in Fig.9.
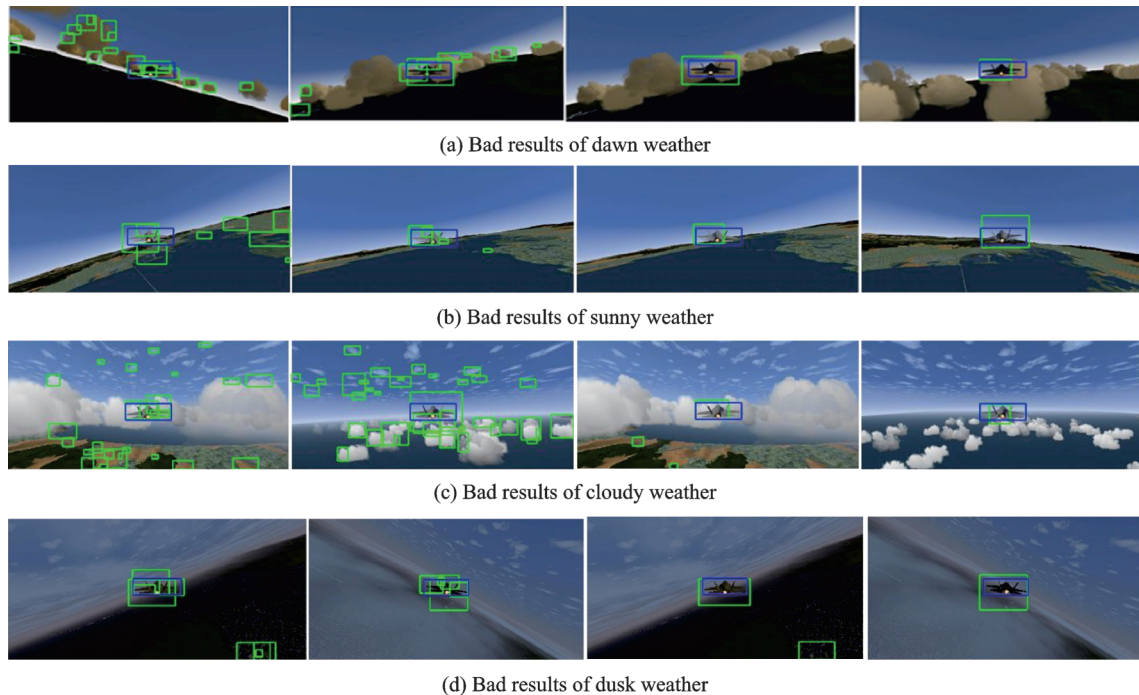
3.2.4   Analysis on results

By analyzing the Figs.4, 5, the detection results in the dawn and the sunny weathers are better than that in the cloudy weather. As the parameters are the best, the recall of the dawn and the sunny weathers is even as high as 0.95, but that of the cloudy weather is only 0.82. In fact, when α and β are increased, the proposals obtained by edge-boxes are also increased until the amount of proposals is 10 000. The parameters are too small to get enough proposals, so the number of ROI boxes is small in the first part of the experiment. The backgrounds of the cloudy weather is more complicated than the others, the light of dusk is less than others, and the color of intruder is similar to the background. These factors would affect the amount of proposals. So the ROI boxes of the cloudy and dusk weathers are usually more. However, this does not mean that more ROI boxes can make the recall higher, instead, the recall may be lower because the detection boxes will be too big to increase the value of IoU after merging the more ROI boxes by NMS method. This can explain why the higher β and α are, the lower the recall is, and can explain why the recalls of cloudy and dusk weathers are lower than the other two.

(a) Good results of dawn weather

(b) Good results of sunny weather

(c) Good results of cloudy weather

(d) Good results of dusk weather

* The blue bounding box is the ground truth, and the green bounding box is the detection result.

Fig.8　Good experimental results of different weathers

(a) Bad results of dawn weather

(b) Bad results of sunny weather

(c) Bad results of cloudy weather

(d) Bad results of dusk weather

* The blue bounding box is the ground truth, and the green bounding box is the detection result.

Fig.9　Bad experimental results of different weathers

By analyzing Figs.6, 7, S and Z can affect the experiments enormously. In the parameter S range of 0.02 to 0.04 and in the parameter Z range of 0.30 to 0.40, the recall is relatively high and the change is gentle. In other words, in the above range, the experimental results are quite good with the average recall more than 0.8. In Fig.5(a), the best recall of cloudy weather only reaches 0.82 with $S=0.03$ and $Z=0.35$. However, in Fig.6(b), the best recall of cloudy weather even reaches 0.9 when S is 0.04. This fact illustrates S is crucial and sensitive in the experiment. In general, we can get a good experi-

ment results by setting $S=0.03$, $Z=0.35$. The two parameters are also changed with a slight step to achieve a better result in the above mentioned range.

In Figs.4—7, it is obvious that the recall rate decreases as the value of IoU increases. When the value of IoU is more than 0.8, the recall rate becomes very small. And when IoU$=$0.5 or IoU$=$0.6, the recall rate is quite high. In other words, IoUs of most detection results is between 0.5 and 0.7, and the detection results are relatively few which have very strict overlap values such as IoU$=$0.8 and IoU$=$0.9.

According to Figs.8, 9, it is visually seen how the ROI boxes affect the final detection results. For those bad results, the main problem is that the detection box is too large or too small to obtain good IoU. In other words, the detection results with low IoU (IoU$\leqslant$0.3) have low location accuracy, and we can regard those detection results as wrong results.

4 Conclusions

A new algorithm is proposed for intruder detection to assist vision-based SAA system. First, this algorithm adopts an optimized edge-boxes method instead of traditional methods to extract the ROIs. Second, sc-SPM is adopted to extract features of test pictures and train the classifier. Finally, a linear SVM classifier is adopted to identify the intruder. Through the optimized edge-boxes method, we can easily obtain a small number of ROIs with good performance. And then the sc-SPM has a very good performance in feature extraction, which can improve the accuracy of the linear SVM classifier. In this paper, a series of experiments have been carried out with different weathers under complex background to discuss the effect of varying parameters on the results and test the effectiveness of the algorithm. The results show that changes in the parameters can greatly affect the experimental results. The results demonstrate our algorithm is effective for intruder detection with different weathers under complex backgrounds, and its performance is more significant in dawn and sunny weathers.

References

[1] YU X, ZHANG Y. Sense and avoid technologies with applications to unmanned aircraft systems: Review and prospects [J]. Progress in Aerospace Sciences, 2015, 74: 152-166.

[2] LYU Y, PAN Q, ZHAO C, et al. A vision based sense and avoid system for small unmanned helicopter [C]//2015 International Conference on Unmanned Aircraft Systems (ICUAS). Denver, USA: IEEE, 2015: 586-592.

[3] ZHANG Z, CAO Y, DING M, et al. An intruder detection algorithm for vision based sense and avoid system [C]//2016 International Conference on Unmanned Aircraft Systems (ICUAS). Arlington, USA: IEEE, 2016: 550-556.

[4] ZHANG Z, CAO Y, DING M, et al. Candidate regions extraction of intruder airplane under complex background for vision-based sense and avoid system [J]. IET Science, Measurement & Technology, 2017, 11(5): 571-580.

[5] CLARK M, KERN Z, PRAZENICA R J. A vision-based proportional navigation guidance law for UAS sense and avoid [C]//AIAA Guidance, Navigation, and Control Conference. San Diego, California, USA:AIAA, 2015.

[6] TAN T, HUANG Y, ZHANG J. Recent progress on object classification and detection[C]//Iberoamerican congress on pattern recognition. Berlin, Germany: Springer, 2013: 1-8.

[7] FELZENSZWALB P F, GIRSHICK R B, Mcallester D, et al. Object detection with discriminatively trained part-based models[J]. IEEE transactions on pattern analysis and machine intelligence, 2010, 32 (9): 1627-1645.

[8] PRIVITERA C M, STARK L W. Algorithms for defining visual regions-of-interest: Comparison with eye fixations[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2000(9): 970-982.

[9] WON C S, SHIRANI S. Size-controllable region-of-interest in scalable image representation [J]. IEEE Transactions on Image Processing, 2011, 20(5): 1273-1280.

[10] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1998 (11): 1254-1259.

[11] MA Y F, ZHANG H J. Contrast-based image attention analysis by using fuzzy growing[C]//Proceedings of the 11th ACM international conference on Multime-

dia. Berkeley, USA：ACM, 2003：374-381.

[12] ZHAI Y, SHAH M. Visual attention detection in video sequences using spatiotemporal cues［C］//Proceedings of the 14th ACM international conference on Multimedia. Santa Barbara, USA：ACM, 2006：815–824.

[13] HAREL J, KOCH C, PERONA P. Graph-based visual saliency［M］. USA：MIT Press, 2007：545-552.

[14] HOU X, ZHANG L. Saliency detection：A spectral residual approach［C］//2007 IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, USA：IEEE, 2007：1-8.

[15] ACHANTA R, ESTRADA F, WILS P, et al. Salient region detection and segmentation［C］//International Conference On Computer Vision Systems. Berlin, Germany：Springer, 2008：66-75.

[16] ACHANTA R, HEMAMI S, ESTRADA F, et al. Frequency-tuned salient region detection［C］//IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2009). Miami, USA：IEEE, 2009：1597-1604.

[17] GOFERMAN S, ZELNIK-MANOR L, TAL A. Context-aware saliency detection［J］. IEEE transactions on Pattern Analysis and Machine Intelligence, 2012, 34(10)：1915-1926.

[18] CHENG M M, MITRA N J, HUANG X, et al. Global contrast based salient region detection［J］. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3)：569-582.

[19] CHENG M M, WARRELL J, LIN W Y, et al. Efficient salient region detection with soft image abstraction［C］//Proceedings of the IEEE International Conference on Computer Vision. Sydney, Australia：IEEE, 2013：1529-1536.

[20] BORJI A, ITTI L. State-of-the-art in visual attention modeling［J］. IEEE transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1)：185-207.

[21] ZHAO Q, KOCH C. Learning saliency-based visual attention：A review［J］. Signal Processing, 2013, 93(6)：1401-1407.

[22] HOU X, ZHANG L. Saliency detection：A spectral residual approach［C］//2007 IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, USA：IEEE, 2007：1-8.

[23] SHEKAR B H, UMA K P, HOLLA K R. Shot boundary detection using correlation based spectral residual saliency map［C］//2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI). Jaipur, India：IEEE, 2016：2242-2247.

[24] SHEKAR B H, UMA K P, HOLLA K R. Shot boundary detection using correlation based spectral residual saliency map［C］//2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI). Jaipur, India：IEEE, 2016：2242-2247.

[25] WANG C, ZHAO L, LIANG S, et al. Object proposal by multi-branch hierarchical segmentation［C］//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA：IEEE, 2015：3873-3881.

[26] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets［J］. Neural Computation, 2006, 18(7)：1527-1554.

[27] BENGIO Y, LAMBLIN P, POPOVICI D, et al. Greedy layer-wise training of deep networks［C］//Advances in Neural Information Processing Systems. Vancouver, Canada：［s.n.］, 2007：153-160.

[28] BENGIO Y, LAMBLIN P, POPOVICI D, et al. Greedy layer-wise training of deep networks［C］//Advances in Neural Information Processing Systems. Vancouver, Canada：［s.n.］, 2007：153-160.

[29] LIU H, LIU S, HUANG T, et al. Infrared spectrum blind deconvolution algorithm via learned dictionaries and sparse representation［J］. Applied Optics, 2016, 55(10)：2813-2818.

Authors　Mr. **ZHANG Zhouyu** received the B.S. and M.S. degrees from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2014 and 2017, respectively. Currently, he is pursuing the Ph. D. degree in Guidance, Navigation and Control at Nanjing University of Aeronautics and Astronautics. His current research interests include Sense and Avoid, object detection and object tracking

Prof. **CAO Yunfeng** received the B. S. and M. S. degrees from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1985 and 1992, respectively. Then he joined the College of Automation Engineering as a Lecturer and became an Associate Professor in 1997. He joined the College of Astronautics in 2003, and became a professor shortly after. He used to be a routine director of Jiangsu Association of Automation, and a vice director of Jiangsu Elec-

tro-technical Society in control section. To date, he has published more than 100 papers in international journals and conferences.

Ms. **ZHONG Peiyi** received the B.S. and M.S. degrees from Nanjing Agricultural University in 2015, and Nanjing University of Aeronautics and Astronautics, in 2018, respectively. Currently, she is an assistant lecturer in Guilin University of Aerospace Technology.

Dr. DING Meng received the Ph.D. degree in Control Theory and Control Engineering from Nanjing University of Aeronautics and Astronautics (NUAA) in 2010. Then he joined the College of Civil Aviation as a Lecturer and became an Associate Professor in 2013. He is a member of the Chinese Society of Astronautics, and frequent reviewer of several international journals. His current research interests include vision-based sense-and-avoid system for UAV, vision-based navigation for UAV, intelligent algorithms for image processing and pattern recognition, and airport surface surveillance techniques. He has undertaken several funded research project as the principle investigator, National Natural Science Foundation of China (2012, 2016), China Postdoctoral Science Foundation (2013), Aeronautical Science Foundation of China (2015) etc. To date, he has published more than 30 papers in international journals and conferences, and proposed three national defense science and technology reports. He was honored the National Defense Science and Technology Progress Award in 2011 (ranking ninth), 2013 (ranking sixth) and 2017 (ranking second).

Mr. HU Yunqiang received the B.S. degree from Nanjing University of Aeronautics and Astronautics in 2012. Currently, he is pursuing the Ph.D. degree in Guidance, Navigation and Control at Nanjing University of Aeronautics and Astronautics.

Author contributions   Mr. **ZHANG Zhouyu** completed the experiments, conducted the analysis and wrote the manuscript. Prof. **CAO Yunfeng** contributed the main idea and the critical advices. Ms. **ZHONG Peiyi** participated in the experiments. Dr. **DING Meng** participated in the discussion. Mr. **HU Yunqiang** participated in manuscript revision.

Competing interests   The authors declare no competing interests.

(Production Editor：Zhang Bei)