

A time-encoded approach to detect cooperative targets for UAV landing

Wang Zi^{1*}, Zhang HongLiang¹, and Shang Yang¹

¹College of Aerospace Science and Engineering, National University of Defense Technology, Changsha 410082, PR China

Abstract. To detect visible light cooperative targets during UAV landing, we use a time relay to encode the intensity of lamps. The change pattern of the lamps intensity is presented in this paper. The frames with strong and weak lamps light are called strong frames weak frames respectively. In continuous sequence images, we take weak frames as background and subtract background using strong frames. The geometric transformations are considered when subtracting background and a float threshold value are used to threshold the difference images. We also design several geometric constraints to eliminate interference points. The approaches to find the first weak frames and to track lamps in a sequence images are also presented. The experiments show that our method is robust to scale change, interference points and moving object in background.

Keywords. UAV landing, cooperative targets, detection.

1 Introduction

The Unmanned Aerial Vehicle (UAV) takes on many military and civilian missions, such as early warning, reconnaissance, surveillance, communication, relay, target indication and damage assessment, and makes it an important tool in modern naval warfare [1]. Because of the complex landing environment and the limited space, UAV landing is one of the most dangerous phases in a flight mission. To solve this problem, several schemes has been proposed, such as net-catch landing, parachute landing, GPS-based landing, IMU-based landing and manual remote control landing [3–5]. Among these proposals, GPS navigation are likely to be disabled in wartime, and it is difficult for other proposals to guarantee security, accuracy, speed and autonomy.

For the last decades, computer vision has been widely used in UAV navigation and landing, due to the development of videometrics and image processing. Vision navigation gets the image of the ideal landing site through airborne camera, detect and recognize the cooperative targets, and obtain the relative position and attitude between UAV and the World Coordinate through coordinate transformation. The accuracy and precision of detection of the cooperative targets are very important in vision-based UAV landing. Although a number of cooperative target schemes have been proposed, infrared lamps are the most popular cooperative targets for fixed wing UAV landing. Gui et al.[2] presented a UAV landing system which consisted of four infrared lamps and a airborne camera with a 940 nm optical filter, and detected the centers of infrared lamps using NLOG operator. In this method, The measurement range of distance between the UAV and the ideal

landing point was about 450 m – 5 m. Hao et al [6] proposed a similar scheme, which consisted a five infrared lamps and a 850 nm optical filter. In his method, the lamp can be detected as far as 80 m. Using infrared lamps and optical filter, cooperative targets detection is robust to illusion, viewpoint, distance and background changing. Actually, an IR filter have to be installed in front of the airborne camera, resulting that the camera is only sensitive to the rays with special wavelength. For most UAVs, it is very hard to add a new camera. So, how to use visible light lamp as cooperate target and how to detect it are valuable engineering problems.

In airborne vision-based landing, visible light lamp detection is different from moving object detection and salient object detection. In moving object detection [7], the background is salient still or dynamic while the object has special motion mode; in salient object detection [8], the object is non-cooperative. Compared with the above two tasks, the cooperative visible lamp detection has three characteristics:

- The lamps are stationary relative to the background.
- The camera is moving, and the scale becomes larger as the UAV approaching the ideal landing site.
- The spatial position of the lamps is fixed, but the temporal intensity can be adjusted.

In this paper, we present a time-encoded visible lamp detection approach in UAV landing. In our method, the intensity of all lamps in continuous video frames changes at a fixed frequency. The video frames with higher intensity of lamp are called the strong frame, while those with lower intensity of lamp are called the weak frame. Because the frequencies of lamps intensity and camera video are fixed, the strong frames and the weak frames

*e-mail: wangzitju@163.com

appear at fixed frequency together. We get background from the weak frames, and remove background in strong frames using difference image. Consider the disturbance points and scale changing in the video during UAV landing, we design several geometric constraints and compensation method. Finally, the effectiveness of the approach is verified by experiments.

2 Structure of Proposed Approach

The equipments used in our approach mainly consist of a camera, four visible light lamps and a digital time relay. The parameters of lamps are shown in Table 1. It should be noted that the visible distance is determined by camera's focal length and aperture. The real lamp is shown in Figure 1. The maximum frequency of time relay is 5 Hz, and the minimum on-off step is 0.1s. The camera we used is Sony FCB-EX2700P, its main parameters are shown in Table 2. The four lamps are installed on an iron shelf, which is shown in Figure 2. The shelf has three arms, which are not in the same plane. The length of the arm towards up is 2.1 m, and the other two are 2.1 m and 3.1 m respectively.

Table 1. The parameters of lamp

power	diameter	emitting angle	visible distance
50 watts	20 cm	30°	300 m

Table 2. The parameters of camera

focal length	aperture	frame rate	video size
3.06 – 122.4 mm	1.6 – 4.6	25 Hz	288 × 352

We use a time relay to blink four lamps synchronously at 5 Hz. The frame rate of the video is 25 Hz. Thus the weak frame should appear every 5 frames. We note \mathbf{F} as the set of a period of continuous video frame, I_i ($1 \leq i \leq M$) as the i_{th} frame, and I_{i_0} as the first weak frame in \mathbf{F} . It should be note that $1 \leq i_0 \leq 5$ and all $I_{i_0+n \times 5}$ ($n \in N$) are weak frames.

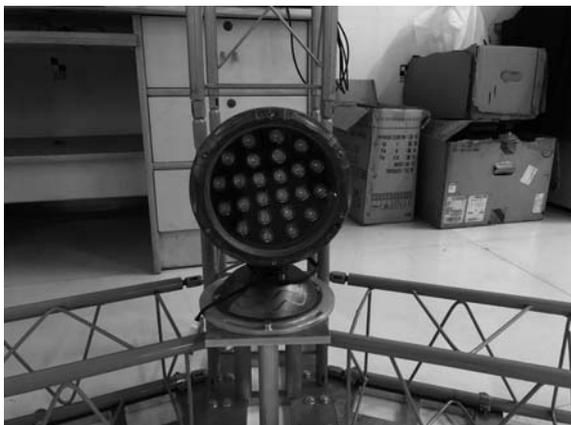


Figure 1. One visible light lamp.

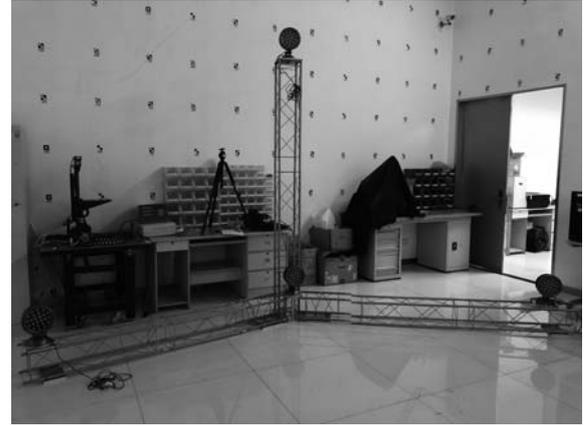


Figure 2. Four visible light lamps with the iron shelf.

The basic idea in this paper is getting background from the weak frames and subtracting the background in the following strong frames to get the position of the lamps. So the first assignment is to find i_0 , in other words, the first weak frame. We assume that the background in adjacent frames are similar. The assumption is tenable, because the camera image the same scene during UAV landing. But the camera is moving, it is impossible that adjacent frames have the same background. We further clarify the assumption that there are translate and scale transformation of background between adjacent frames. The reasons why we not use affine and projective transformation are

- affine and projective transformation have too many parameters;
- it is difficult to get a lot of stable feature points during UAV landing, because of the spurious feature points of ideal landing site;
- the dissimilarity between adjacent frames is relative small as long as UAV's attitude doesn't change drastically.

The second assignment is the estimation of the transformation from weak frame I_{i_0+5n} to the following four strong frame I_{i_0+5n+j} ($j = 1, 2, 3, 4$) respectively. And the transformation matrix is noted as \mathbf{S}_n^j , where $n = 0, 1, 2, 3 \dots$ and $j = 1, 2, 3, 4$. Then the transformed weak frames I'_{i_0+5n} can be obtained using geometric transformation:

$$I'_{i_0+5n} = \mathbf{T}(I_{i_0+5n}, \mathbf{S}_n^j) \quad (1)$$

where $\mathbf{T}(\cdot)$ is geometric transformation. And the background can be removed from the strong frames using difference image:

$$D_{i_0+5n+j} = I_{i_0+5n+j} - I'_{i_0+5n} \quad (j = 1, 2, 3, 4) \quad (2)$$

where D_i is the difference of I_i and the I'_{i-1} . Then we threshold the difference image, extract 8-connect regions and calculate their centers. The binary image of D_i is noted as B_i . We take all the centers as the candidates of the position of center. Finally, several geometric constraints of the relative position relation between lamps are presented, to

determine lamp position from the candidates. We adopt a floating threshold when thresholding the difference image:

- if there are too many candidates, we increase the increase threshold value.
- if there is no combination of four centers satisfying the constraints, we decrease the threshold value.

The work flow of the proposed method is shown in Figure 3.

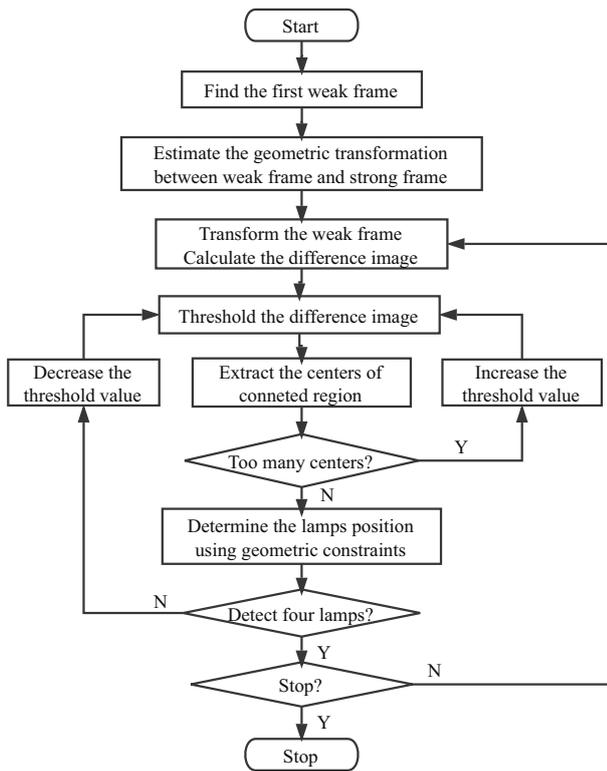


Figure 3. The work flow chart of the visible light lamp detection.

3 Principle and technical implementation

Before mentioned, several assignments have to be tackled to implement the proposed method. In this chapter, we present the approaches for these problems.

3.1 Estimation of transformation

We take each weak frame as the background of the next four strong frames. Because of the motion of the camera, the transformation relationships between weak and strong frames need to be estimated. During UAV landing, only the image of the scene around the ideal landing site or the lamps are stable, due to dynamic background. Naturally, we should only use the image of the ideal landing site or scene around the lamps to estimate the transformation.

For a long time, SIFT+BFMacth+RANSAC has been the most used method for estimation of geometric transformation between image pairs, which is a general approach definitely. Recently, a new method for feature matching in

video application, Grid-based Motion Statistics(GMS)[9], has been proposed. GMS uses ORB features and provides a real-time, ultra-robust correspondence system. We use ORB+BFMatch+GMS to detect and match the feature points in weak and strong frames. We note the set of matched feature points in weak and strong frames as $\{p_i = (x_i, y_i, 1)^T\}$ and $\{p'_i = (x'_i, y'_i, 1)^T\}$ respectively. The transform matrix from weak frame to strong frame is given as:

$$\mathbf{S} = \begin{bmatrix} s & 0 & d_x \\ 0 & s & d_y \end{bmatrix} \quad (3)$$

where s is scale factor while d_x and d_y are translation factor. So, we get the equation:

$$p'_i = \mathbf{S}p_i \quad (4)$$

for N pairs matched points, we have $2N$ equations actually:

$$\begin{bmatrix} x'_1 \\ y'_1 \\ x'_2 \\ y'_2 \\ \vdots \\ x'_N \\ y'_N \end{bmatrix} = \begin{bmatrix} x_1 & 1 & 0 \\ y_1 & 0 & 1 \\ x_2 & 1 & 0 \\ y_2 & 0 & 1 \\ \vdots & \vdots & \vdots \\ x_N & 1 & 0 \\ y_N & 0 & 1 \end{bmatrix} \begin{bmatrix} s \\ d_x \\ d_y \end{bmatrix} \quad (5)$$

the equation (5) can be easily solved by Least Square method.

3.2 Geometric constraints

If we find N connected region centers and $N \geq 4$, C_N^4 candidates of the position of four lamps need to be examined. We design several geometric constraints according to the relative position of lamps. Figure 4 is the binary image of the difference image.

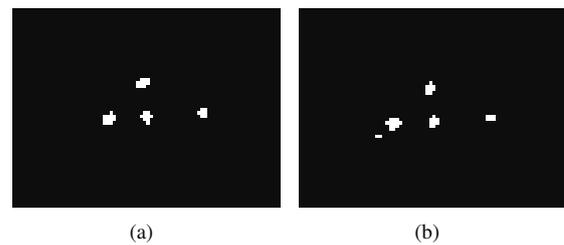


Figure 4. The binary image of difference image. (a): without interference point; (b): with interference point.

To examine four points in a candidates, the first step is numbering them as Table 3. To illustrate our approach effectively, we note the position of the coordinates of the numbering centers as (x_U, y_U) , (x_L, y_L) , (x_M, y_M) and (x_R, y_R) respectively. The second step is designing several relative position constraints:

1. $x_L \leq x_U \leq x_R$
2. $1/th_1 \leq Dis(U, M)/Dis(L, M) \leq th_1$

3. $|(U - M)^T(L - M) / (Dis(U, M) \cdot Dis(L, M))| \leq th_2$
4. $(L - M)^T(R - M) / (Dis(L, M) \cdot Dis(R, M)) \leq th_3$
5. $th_4 \leq Dis(R, M) / Dis(L, M) \leq th_5$

The $Dis(\cdot)$ above is L_2 distance. The constraint 1 means that x coordinate the top lamp should be smaller than the right one's and be bigger than the left one's; the constraint 2 and 3 mean that the three centers on the left are similar to an equilateral triangle; the constraints 4 means that the three centers on the bottom are in a line; the constraint 5 means that the ratio the distances from the left one and the right one to the middle one are approximately equal to 1.

Table 3. The approach for numbering centers

number	position
U	the center with the smallest y
L	the center with the smallest x expect U
R	the center with the biggest x expect U and L
M	the remaining one

3.3 Track centers of lamps in sequence images

If four centers coordinates satisfy the constraints above, we use a similar method in [2] to track lamps in sequence images. But the different thing is we only detect lamps in strong frames. As four centers in a strong frame satisfy the constraints above, the maximum distance D_{xmax} and D_{ymax} in x and y axis among four centers and their average coordinates is computed. The average coordinates are computed as:

$$\begin{cases} \bar{x} = (x_U + x_L + x_M + x_R) / 4 \\ \bar{y} = (y_U + y_L + y_M + y_R) / 4 \end{cases} \quad (6)$$

In the next strong frame, the searching is done in the region which takes (\bar{x}, \bar{y}) as the center and $3D_{xmax}$ and $3D_{ymax}$ as width and height respectively. However, it is difficult to detect and match feature points in Section 3.1 when the searching area is very small. So we choose a threshold for the width and height of searching region. We donate the width and height as W and H respectively, then they are given as:

$$W = \begin{cases} 3D_{xmax}, & \text{if } 3D_{xmax} \geq 80 \\ 80, & \text{otherwise.} \end{cases} \quad (7)$$

$$H = \begin{cases} 3D_{ymax}, & \text{if } 3D_{ymax} \geq 60 \\ 60, & \text{otherwise.} \end{cases} \quad (8)$$

Because of the shake of the UAV, the scene of the lamps is likely to be out of the camera during UAV landing. In this case, if no centers satisfy the constraints in continuous four strong frames, the searching region are set to the whole image otherwise the searching region is not updated.

3.4 Further details

We stated earlier that the first assignment is find the first weak frame. Because two adjacent of a weak frame are strong frames, there should be four centers satisfy the constraints in the binary images of the difference images of a weak frame and the previous frame and the next frame. In other words, if I_i is a weak frame, I_{i-1} and I_{i+1} are strong frames. Therefore, there must four centers satisfy the constraints above in binary image B_i and B_{i+1} . So our approach to detect the first weak is find i_0 so that B_{i_0} and B_{i_0+1} contain four centers satisfy the constraints and the searching region is the whole image.

After thresholding the difference image, some morphological operations are taken to improve the robustness of our approach. The first operation is eliminating connected region whose area is smaller than 4. Then morphological dilation is taken to connect disconnected regions caused by subtracting images. Finally, the connected regions are counted and each center is computed as:

$$p_{center} = \frac{1}{Area} \sum_{p_i \in region} p_i \quad (9)$$

Once i_0 is obtained, the positions of four lamps are also obtained and the searching region for all the following strong frames can be reduced as described in Section 3.3.

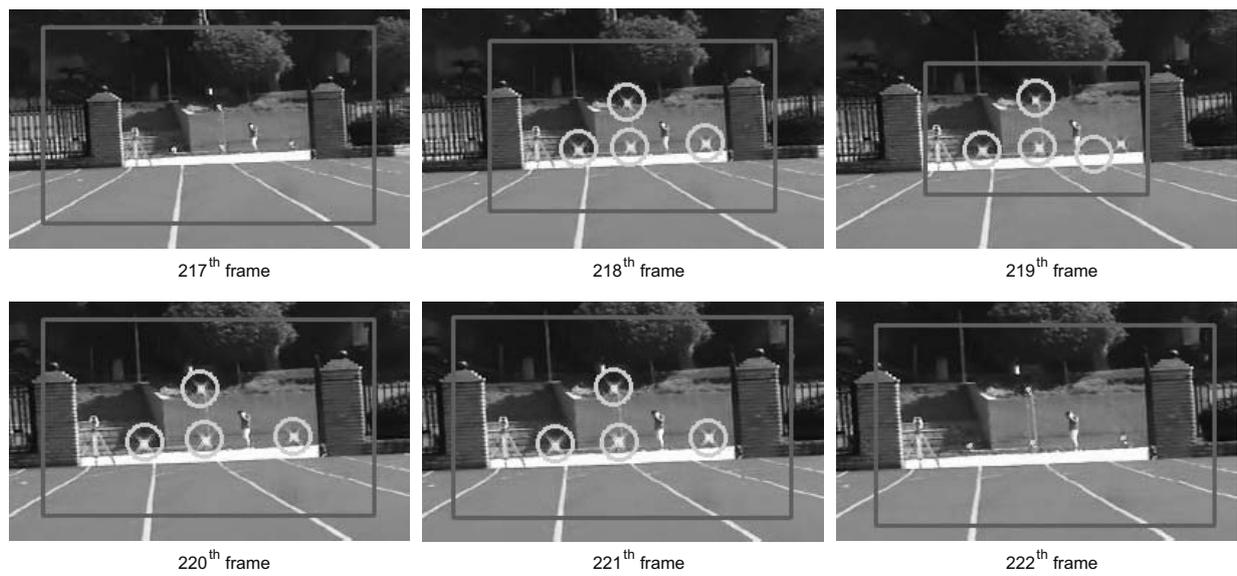
4 Experimental results and discussion

To prove the validity of the method, experiments were carried out. The videos were taken by the equipments mentioned in Section 2. We did not optimize any component of our final method. The program was written by OpenCV 2.4.13 using C++ and run on Intel 7700k at 4.2 Hz with 16 GB RAM. The distance from lamps to the camera is about 200 m, measured by a total station. The parameters used in Section 3.2 are shown in Table 4, which is determined by experiments.

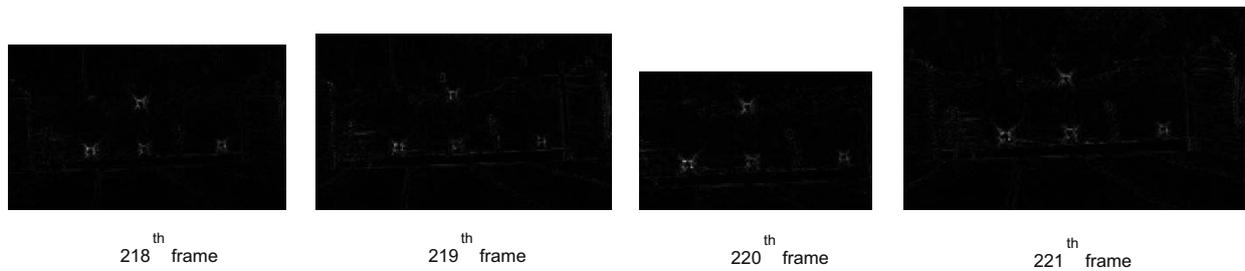
Table 4. The parameters of geometric constraints

th_1	th_2	th_3	th_4	th_5
1.5	0.2	-0.95	1.0	2.0

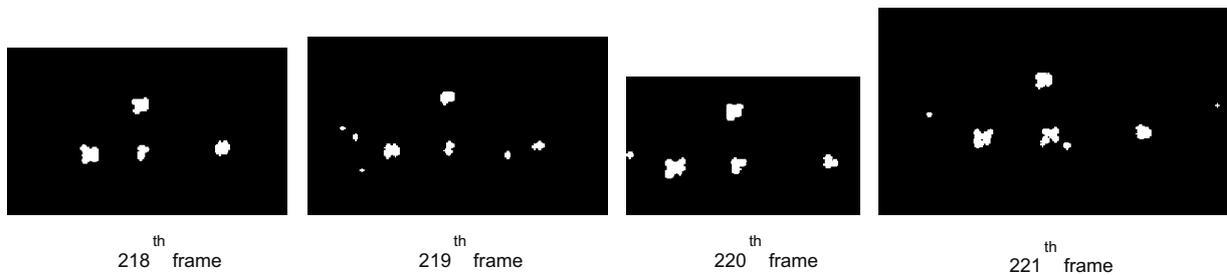
The result of our approach is shown in Figure 5. The original image size is 288×352 and the result images in Figure 5 and 6 are cropped in order to demonstrate our method effectively. The intensity of the lamp changes with a fixed mode, which can be seen visually. The 217th and 222th frame are weak frames while the other four among them are strong frames. The scale change can be seen visually, caused by the change of focal length, as the person and the shrub become bigger. To illustrate the robust of our method, the difference images and binary images are also shown. Unfortunately, the detection of the right lamp was failure in 219th frame, because the interference point and the other three lamps in the the binary image satisfy the constraints in Section 3.2. However, this failure only effect the next searching region's position and



(a) the detection result(the images are cropped)



(b) the difference image between strong frames and the weak frame



(c) the binary images of difference images

Figure 5. The detection of lamps in scale changed sequence images.

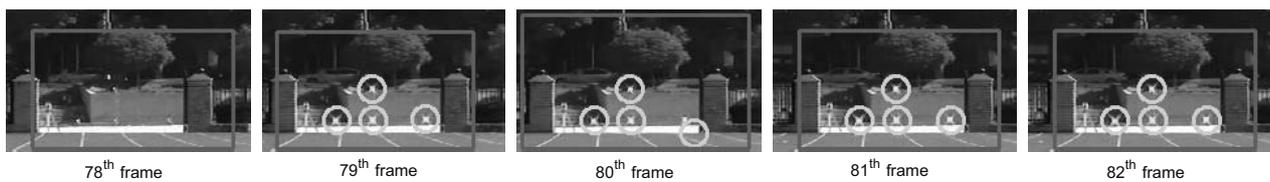


Figure 6. The detection of lamps in sequence images with a white car going through searching regions(the images are cropped).

size. And the success of 220th and 221th frame show that our method is robust to such failure. In 220th and 221th frame, the binary images also show the effectiveness of the geometric constraints in Section 3.2. In Figure 6, a white car went through four searching regions, but the lamps were detected successfully, which show the method is robust to such interference. The computation time of one frame ranges from 40 ms to 100 ms, determined by the size of searching region. The average computation time is about 70ms for one frame, which satisfy the processing frequency of the airborne control system proposed in [2].

5 Conclusion

To utilize visible light lamps as the cooperative targets for UAV landing, we use a time delay to change the intensity of lamps temporally and analysis the pattern of change. In our experiments, there is an image in which the intensity of lamps is weak among five continuous sequence images, while the intensity of lamps in the other four image is strong. The former type frames are called weak frames and the latter ones are called strong frames. We take weak frames as background and subtract background from strong frames. The geometric transformations between weak and strong frames are considered and a float threshold value strategy is used. In order to find lamps, we design a approach to eliminate interference points. The method for tracking lamps and finding the first weak frame in sequence images are presented in this paper.

The experiments show that the proposed method is robust to moving object in background, scale change and interference points. And the computation time is acceptable. In conclusion, our method is suitable for detecting visible light cooperative targets during UAV landing. The further work can be done on the location accuracy of lamps and the constraints to eliminate interference points.

Acknowledgement: This work was supported by Scientific Research Program of National University of Defense Technology(No.ZK16-03-27).

References

1. Kumar V, Yong H, Min D, et al. Auto landing control for small scale unmanned helicopter with flight gear and HILS. International Conference on Computer Sciences and Convergence Information Technology. IEEE, 2011:676-681.
2. Gui Y, Guo P, Zhang H, et al. Airborne Vision-Based Navigation Method for UAV Accuracy Landing Using Infrared Lamps. Journal of Intelligent & Robotic Systems, 2013, 72(2):197-218.
3. Coutard L, Chaumette F. Visual detection and 3D model-based tracking for landing on an aircraft carrier. IEEE International Conference on Robotics and Automation. IEEE, 2011:1746-1751.
4. Zhang Y, Zhao S S, Wang Y S, et al. Study on Guidance Law of UAV Net Recovery System. Journal of China Ordnance, 2007, 3(1):64-67.
5. Guo L, Zhang H Y, Tong M B. Dynamic analysis on parachute recovery of aerial vehicle. Journal of Nanjing University of Aeronautics & Astronautics, 2012, 44(1):14-19.
6. Hao S, Cheng Y, et al. Robust detection and matching algorithm oriented to near infrared cooperative objects. System Engineering and Electronics, 2014, 36(9):1854-1859.
7. Shantaiya S, Verma K, Mehta K. A Survey on Approaches of Object Detection.pdf. IEEE Electron Device Letters, 2013, 28(5):419-421.
8. Borji A, Cheng M M, Jiang H, et al. Salient Object Detection: A Survey. Eprint Arxiv, 2014, 16(7):3118.
9. Bian J W, Lin W Y, Matsushita Y, et al. Gms: Grid-based motion statistics for fast, ultra-robust feature correspondence. IEEE Conference on Computer Vision and Pattern Recognition. 2017.