Research on distance measurement method of moving objects in video

Wei Zhou¹, Youfeng Zhang¹, Yunhong Zheng¹*, Pingfan Li¹, Min Xu¹
Xiamen University of Technology, Xiamen, China

*Corresponding author e-mail: zhengyunhong@xmut.edu.cn

Abstract. We propose a smart car loading camera to simulate the vehicle driving process, the purpose of which is using monocular vision to solve the collision problem that may occur when passengers open the door after the vehicle stops. First of all, Youzheng Zhang's calibration method is used to obtain the internal and external parameters of the camera, and then the inter-frame difference method is used to lock the moving target in the video. After preprocessing the image, an iterative method is proposed to automatically obtain the threshold to complete the target object segmentation. Combined with the calibration parameters, the distance of the target object is calculated based on the proportion of similar triangles. The test results of videos taken in multiple different scenes show that the average measurement error of the algorithm is 44mm.

1. Introduction

When the driver or passenger of the car opens the door and gets off the vehicle, if they do not notice the moving vehicle or pedestrian behind, accidents are likely to occur. In fact, accidents caused by the door opening often occur, which may lead to serious casualties. This article uses a monocular camera to shoot a video of the rear scene to determine whether it is possible for a moving object in the rear to hit the door in the open state.

The moving object detection method realizes the detection of target objects through the existence of a large amount of strong correlation information between adjacent frames of a video stream sequence. At present, the optical flow method and the difference method are more widely studied, and can be divided into the difference between the background and the video frame according to the difference object. Machine vision ranging uses a camera to convert three-dimensional spatial position information into two-dimensional image information. The target position information in the two-dimensional image is calculated by a mapping model to measure the position information of the target in the three-dimensional world. At present, there are mainly two types of ranging methods: monocular ranging and binocular ranging. Jie Dong proposed a defocus ranging algorithm based on monocular vision. Using the two-dimensional gray level information of the image, it is difficult to determine the positional relationship between the focused image plane and the specific imaging plane in the single image defocus ranging algorithm by using the gray gradient method. Jun Wu, Wenjie Li, et al combined multi-scale block binarization mode and Adaboost method to extract candidate regions of vehicles. Then, the interference is eliminated according to the horizontal edge features and gray features in the extracted area. Finally, the distance information of the vehicle was measured by an improved positioning method based on the shadow feature of the bottom of the vehicle.
2. Target Object Detection Based on Threshold Segmentation

In this paper, the background difference method is used to extract the moving object image in the video, and then the iterative method is used to automatically obtain the threshold value for image segmentation. Finally, according to the characteristics of the target object, the minimum circumscribed rectangle is used to obtain the center of mass coordinates of the target object.

2.1. Moving object detection in video

The background difference method is to extract the moving target by subtracting the image of the current frame from the background image established in advance. The basic steps are:

a) Acquire the background model image $f_b(x, y)$.

b) Differentiate each current frame image $f_k(x, y)$ from the background image $f_b(x, y)$ to obtain the foreground image.

c) Finally, using a method similar to inter-frame difference, the two parts are cycled to obtain the foreground image of each frame of the moving object.

The specific formula is as follows:

$$D_k(x, y) = \begin{cases} 1, & |f_k(x, y) - f_b(x, y)| > T \\ 0, & \text{other} \end{cases}$$

In this paper, the median filtering method and dilation algorithm are used to preprocess the extracted foreground image to eliminate light pollution and noise pollution caused by slight vibration, and enhance the robustness of the algorithm.

2.2. Iterative method to obtain the threshold

The principle of the iterative method is: first, the user selects an initial threshold, then continuously tests and changes through the algorithm, and finally determines the target threshold size. The specific process is as follows:

a) In the grayscale picture of [0, L] to be used, the user selects a threshold $T$ as the initial threshold, and generally selects the median of the grayscale;

b) The user uses the threshold to divide the picture into two parts, and calculates their average threshold using the following formula:

$$A = \frac{\sum_{i=1}^{T_l} I_i n_i}{\sum_{i=1}^{T_l} n_i}, \quad B = \frac{\sum_{i=T_l+1}^{L} I_i n_i}{\sum_{i=T_l+1}^{L} n_i}$$

c) Bring results A and B into a new threshold $T_{i+1}$;

$$T_{i+1} = \frac{1}{2} \times (A + B)$$

d) Repeat steps 2 and 3 until $T_{i+1}$ and $T_i$ is less than the specified value.

2.3. Minimum circumscribed rectangle to get the target centroid coordinates

Moving target feature locking methods include shadow method, optical flow method, trajectory method, etc. The main method adopted in this paper is to select two features of histogram and centroid to lock the target.

The background difference method is used to obtain the area of the moving target, and then all the areas of the moving target are analyzed, and the data of the area such as $X_{\text{max}}$ and $X_{\text{min}}$ of the $X$ axis; $Y_{\text{max}}$ and $Y_{\text{min}}$ of the $Y$ axis are then obtained, and then the four algorithms are connected using matlab's algorithm Point, you can get the smallest circumscribed matrix of the target object.

The height of the smallest circumscribed matrix is:
3. Distance measurement based on monocular vision

In this article, Zhengyou Zhang’s calibration method is used to obtain the parameters of the camera model, and the similar triangle method is used to calculate the horizontal distance from the moving object to the camera.

3.1. Camera calibration

Camera calibration refers to the relationship between the pixel position of the camera image and the position of the scene point. According to the pinhole camera principle, the parameters of the camera model are solved from the correspondence between the coordinates of the feature points in the image and the world coordinates. The model parameters that the camera needs to calibrate include internal parameters and external parameters. The pinhole imaging model is shown below.

\[
H = X_{\text{max}} - X_{\text{min}}
\]

The width is:

\[
L = Y_{\text{max}} - Y_{\text{min}}
\]

The coordinates of the center of mass:

\[
\left( \frac{X_{\text{min}} + H}{2}, \frac{Y_{\text{min}} + L}{2} \right)
\]

3.2. Distance measurement

The distance measurement uses the similar triangle method. According to the proportional relationship between the pixel coordinate system, the image coordinate system, and the world coordinate system, a mathematical model is established through mathematical knowledge, as shown in the figure below.

![Figure 1. Pinhole imaging model.](image-url)
Figure 2. Similar triangle measurement model

In the figure 2, the camera lens height $H$, the distance between the lens target point corresponding to the pixel center and the camera $O_3M$, the focal length $f$, and the scale factors $fx, fy$ are known. The calculation process is as follows:

$$\alpha = \arctan\left(\frac{H}{O_3M}\right)$$

(7)

$$\gamma = \arctan\left(\frac{O_3P_1 \cdot x\text{pix}}{f}\right)$$

(8)

$$\beta = \alpha - \gamma$$

(9)

Get the vertical distance:

$$O_3P = \frac{H}{\tan B}$$

(10)

$$O_{1P_1} = \sqrt{\left((v - v\text{center}) \cdot x\text{pix}\right)^2 + f^2}$$

(11)

$$O_{2P} = \frac{H}{\sin B}$$

(12)

From $\frac{PQ}{P_2Q_2} = \frac{O_2P}{O_2P_1}$

Receive $PQ = \frac{O_2P_2 \cdot P_2Q_2}{O_2P_1}$

(13)

4. Test and result analysis

In order to evaluate the accuracy of the article's algorithm, in a natural light environment, a Huawei mobile phone equipped with a Meconium wheel cart was used as the hardware platform, and the algorithm was verified on MATLAB. There were 8 groups of video shots, and each group took 50 frames.

4.1. Verification of distance measurement algorithm for moving objects

Now select a group from the above images for detailed algorithmic process analysis. The target object detection results are shown in Figure 3. Figure 3a is the original image obtained after the image is framed. The binary image obtained after image preprocessing and threshold segmentation is shown in Figure 3b, and Figure 3c is the position information of the target object obtained using the smallest circumscribed rectangle.
c. Get the centroid coordinates of the smallest circumscribed rectangle

Figure 3. Target object detection based on threshold segmentation

This article uses the calibration toolbox in the matlab toolbox to calibrate the camera. The camera takes six checkerboard photos. The process is shown in Figure 4.

Figure 4. Camera calibration

Get camera calibration data: \( f_x = 426 \text{pix}, f_y = 429 \text{pix}, u0 = 372 \text{pix}, v0 = 146 \text{pix} \). According to the formula \( f_x = f / dy, f_y = f / dy, dx \approx dy = 0.00939 \) can be measured. Bring the data into the formula to get the distance value.

4.2. Algorithm evaluation and error analysis

In this paper, distance measurement is performed on multiple points in the image, and the measurement results of 4 points are listed.

Table 1. Algorithm error.

| Numble | Test point | Image coordinates | Actual distance from moving object (mm) | Target distance calculation (mm) | difference (mm) |
|--------|------------|-------------------|----------------------------------------|---------------------------------|----------------|
| 1      | Left two   | (0, 98)           | 1149                                   | 1104.23                         | 44.77          |
| 2      | Left four  | (165, 98)         | 1110                                   | 1066.43                         | 43.57          |
| 3      | Left six   | (283, 98)         | 1006                                   | 995.9                           | 50.10          |
| 4      | Mid point  | (372, 98)         | 938                                    | 975.24                          | 37.24          |
This test counts the measurement results of 100 points. It can be obtained that the average distance measurement error rate of the article algorithm is 4.4mm, which can basically meet the accuracy requirements of ranging.

5. Conclusions
The article studies the distance measurement method of the target object in the video, and the conclusions are as follows:

a) In this paper, the background difference method is used to extract the moving object image in the video, then the iterative method is used to automatically obtain the threshold value for image segmentation, and finally the minimum outer rectangle is used to obtain the image coordinate value of the target object.

b) This article uses Zhengyou Zhang’s calibration method to obtain the parameters of the camera model, and then uses the similar triangle method to calculate the horizontal distance from the moving object to the camera.

c) Under the natural light environment, the average error of the algorithm in this paper is 44mm. The analysis results show that the algorithm is robust and the accuracy of the algorithm needs to be improved.

6. Acknowledgments
Project Supported by Education research project for young and middle-aged teachers of Fujian (JT180428).

References
[1] Linhao Li, Qinghua Hu, Xin Li. Moving Object Detection in Video via Hierarchical Modeling and Alternating Optimization, [J]. IEEE transactions on image processing : a publication of the IEEE Signal Processing Society, 2018.
[2] Zhiming Qian, Shuohong Wang, Xi En Cheng et al. An effective and robust method for tracking multiple fish in video image based on fish head detection, [J]. BMC Bioinformatics, 2016, 17(1).
[3] Daofu Chen. Research on Video Object Retrieval Technology Based on Deep Learning [D]. University of Electronic Science and Technology of China, 2019.
[4] Yarahmadi Reza, Bagherpour Raheb, Kakaie Reza, Mirzaie N. Hossein, Yari Mojtaba. Development of 2D computer program to determine geometry of rock mass blocks [J]. International Journal of Mining Science and Technology, 2014, 24(02):191-194.
[5] Lu ping Wang, Lu ping Zhang, Ming Zhao, Biao Li. An algorithm for moving target detection in IR image based on grayscale distribution and kernel function [J]. Journal of Central South University, 2014, 21(11):4270-4278.
[6] Dongxue Sun. Video Object Segmentation to Enhance Connectivity and Consistency of Internal Areas of Targets, [D]. Chongqing University, 2018.
[7] Wenminghu, Jianguan, You He. Distance extended target detection based on video accumulation, [J]. Fire Control Radar Technology, 2006 (04): 16-18 + 22.
[8] Chunyu Zhang. Research on Video Object Tracking Algorithm Based on Motion Features, [D]. Shenzhen University, 2016.