Target motion detection algorithm based on dynamic threshold

Zhao Zeyi\textsuperscript{1}, Lu Gang\textsuperscript{2}\textsuperscript{*}
\textsuperscript{1}Finance Office, Shaanxi Normal University, Xi’an, 710119, China
\textsuperscript{2}School of Computer Science, Shaanxi Normal University, Xi’an, 710119, China
\textsuperscript{*}Corresponding author’s e-mail: lg@snnu.edu.cn

Abstract. Realizing moving target detection through visual algorithms is a major branch of computer technology. Moving target detection has a wide range of applications in many fields such as video surveillance, unmanned driving, and behavior research. The current algorithm has been able to achieve moving target detection well in a static background, but the detection of multiple moving targets in a dynamic background is still a challenging research area. The main pain points are complex background, light changes, noise interference and Occlusion and real-time detection requirements, etc. Aiming at the detection of moving targets in complex scenes, a moving target detection algorithm based on dynamic threshold is proposed. The algorithm first uses the gray difference gradient of three adjacent frames of the video sequence to calculate the segmentation threshold of foreground moving targets to realize the dynamic threshold. Then use dynamic thresholds to achieve multi-target motion detection in complex backgrounds or scenes with changing lighting. Because the gradient change of adjacent frames of the scene is considered in the process of moving target detection and segmentation, and the gradient change is converted into a dynamic threshold to participate in moving target detection and segmentation, the algorithm can adapt to scenes with changing lighting. Experiments show that this algorithm is feasible for moving target detection in complex scenes, and has a lower error rate than that of traditional algorithms, and is more suitable for to complex or changing lighting scenes.

1. Introduction

Moving target detection has a wide range of applications in life, such as environment monitoring, unmanned driving, behavior analysis and other fields. The traditional target motion detection algorithm mainly builds a pixel model for the background, and solves the area of the foreground moving target through the difference between the current frame and the background model. For example, the moving target detection algorithm based on the Gaussian mixture model is proposed and applied\textsuperscript{[1-3]}. Literature\textsuperscript{[1]} proposed a method of using the Gaussian mixture method to model the background. By modeling and continuously updating the background of the real scene, the Gaussian mixture method can effectively overcome the tiny disturbance noise caused by camera shake. The Gaussian mixture method is widely used to detect moving targets in video sequences\textsuperscript{[2-3]}, but the Gaussian mixture method has the disadvantage of being sensitive to rapid illumination changes. Most scenes in life change over time. For example, there is camera shake to capture the scene, weather changes, occlusion and lighting changes cause the captured scene to have a dynamic background, which cause more difficult to detect moving objects. Literature\textsuperscript{[4]} proposed a moving target detection method combining the continuous frame difference method and the background difference method.
The algorithm uses linear adaptive filtering and nonlinear median filtering to differentiate the obtained background image. At the same time, the threshold segmentation technology is used to achieve the enhancement of the moving target, so as to effectively solve the phenomenon that the target can not be detected in the background difference method and the inter-frame difference method. Literature[5] proposed a moving target detection algorithm that fused difference method and vibe algorithm. The algorithm realizes the elimination of ghost image pixels by judging the temporal changes of two consecutive frames of foreground and background pixels. Literature[6] proposed a moving target detection algorithm based on the smooth three-frame difference method fusion RPCA. The algorithm uses the RPCA algorithm to extract the background of the current frame as the intermediate frame of the smooth frame difference method, thereby eliminating the "hole" phenomenon and greatly reducing noise. Literature[7] proposed a moving target detection algorithm based on pixel spatial sample difference consensus. The algorithm compares the values of two pixels occupying the same position in adjacent frames to determine the moving object area, sets a spatial sample set for each pixel, and defines the spatial sample difference consensus (SSDC), which represents the change of stable spatial relationship. By calculating the SSDC between two adjacent frames to subtract the moving object, so as to achieve a better moving target detection effect. Literature[8-10] show the practical application of moving target detection algorithm in dynamic background. This paper presents an adaptive threshold multi-target motion detection algorithm. By introducing the gradient change of the pixel difference between adjacent frames into the segmentation threshold calculation of moving target detection, the algorithm can adapt to the changes of the environment, camera shake and the lighting changes of the scene.

2. Algorithm

In computer vision, the main difficulty of moving target detection algorithms is how to eliminate noise interference in the scene, such as camera shake, light changes, environmental changes. In complex scenes, it is a challenging task to distinguish the foreground moving target from the background through a single threshold. The algorithm proposed in this paper participates in the calculation of the threshold through the gradient of the pixel difference between adjacent frames, and divides the moving target from the background through the dynamic threshold calculation to achieve the target motion detection with a much smaller error rate.

2.1. Background model generation

When the algorithm is initialized, it is necessary to first convert the input frame into a gray image Grayb, and solve its corresponding binary image BGrayb. When there is no moving target in the scene or the moving target is not moving in a short period of time, the gradient of the gray level difference between adjacent frames is close to zero. The position of the moving target cannot be determined by gradient changes, but this problem can be solved by constructing a static background model and allowing it to participate in dynamic threshold calculation.

2.2. The gray difference of adjacent 3 frames

Let Num_Frame be the number of video frames, the algorithm starts from the second frame to obtain the gray-scale pictures of each frame in turn, and formula (1) represents the gray-scale pictures of all frames.

\[
\text{Gray}_f = \text{FrameRGB}_\text{gray}(f), \quad \text{Gray}_f = \{\text{gray}(i, j) | 1 \leq i \leq \text{row}, 1 \leq j \leq \text{col}\}
\]  

(1)

Among them, Gray\(_f\) is the grayscale picture of the current frame \(f\), row is the number of rows in the picture, and col is the number of columns in the picture.

The gray level difference of the adjacent 3 frames is del1, del2, del3, and the method of obtaining it is formula (2)
\[ \begin{align*}
\text{del1} &= |\text{Gray}_{t,1} - \text{Gray}_t | \\
\text{del2} &= |\text{Gray}_{t+1} - \text{Gray}_t | \\
\text{del3} &= |\text{Gray}_b - \text{Gray}_t | \\
\end{align*} \tag{2} \]

2.3. Dynamic threshold solution

Participate in the dynamic threshold calculation of the gray difference gradient value of adjacent frames, and solve the dynamic threshold \( \text{Threshold}_{\text{image}} = \text{Threshold}(\text{del1, del2}) \)

2.3.1 Step 1, Solve &

Before obtaining the dynamic threshold, the algorithm in this paper will use the current frame to initialize an initial threshold. The method of obtaining it is shown in formula (3)

\[ \& = \frac{1}{\text{row} \times \text{col}} \left( \sum_{\text{col}} \max(\text{del}_1) + \sum_{\text{col}} \max(\text{del}_1^T) \right) \]

(3)

Among them, del1 is the gray level difference between the current frame and the previous frame, \( \text{row} \) is the number of rows in the picture, and \( \text{col} \) is the number of columns in the picture.

2.3.2 Solve the Threshold_image of adjacent frames

If \& is not equal to 0, there are moving targets in the scene at this time. Solve the Threshold_image of adjacent frames by formula (4) and formula (5).

\[ \text{Threshold}_{\text{image}} = \{ V(i, j) \mid 1 \leq i \leq \text{row}, 1 \leq j \leq \text{col} \} \]

(4)

\[ V(i, j) = \begin{cases} 
0, & \text{del1}(i, j) \leq \& \\
1, & \text{del2}(i, j) \geq \& 
\end{cases} \]

(5)

Among them, del1 is the gray level difference between the current frame and the previous frame, del2 is the gray level difference between the current frame and the next frame, \( \text{row} \) is the number of rows in the picture, and \( \text{col} \) is the number of columns in the picture.

If \& is equal to 0, there is no moving target in the scene at this time, or the moving target has not moved in a short time. Solve the Threshold_image of adjacent frames by formula (6) and formula (7).

\[ \text{Threshold}_{\text{image}} = \{ V(i, j) \mid 1 \leq i \leq \text{row}, 1 \leq j \leq \text{col} \} \]

(6)

\[ V(i, j) = \text{del2}(i, j) \]

(7)

Among them, del2 is the gray level difference between the current frame and the next frame, \( \text{row} \) is the number of rows in the picture, and \( \text{col} \) is the number of columns in the picture.

2.3.3 Solve the dynamic threshold of the current frame

Solving the dynamic threshold of the current frame is divided into 4 steps. Firstly, the dynamic threshold of the current frame and the background model is solved; secondly, the result of step 1 is combined with the gray difference of the previous frame to optimize the dynamic threshold; third, the result of step 1 is combined with the gray difference of the next frame to further optimize the dynamic threshold. Finally, the maximum value of the results of steps 1-3 is used as the dynamic threshold of the current frame. See equations (8), (9), (10), (11) for specific solutions.

\[ \text{Th}_1 = \text{Threshold}(\text{del1, Bgray}_b) \]

(8)

\[ \text{Th}_2 = \text{Threshold}(\text{del3, Bgray}_b) \]

(9)

\[ \text{Th}_3 = \text{Threshold}(\text{del2, Th}_1) \]

(10)

\[ \text{Result}_j = \max(\text{Th}_1, \text{Th}_2, \text{Th}_3) \]

(11)
Among them, $B_{grayb}$ is the binarized picture of the background picture, $delt1$ is the gray level difference between the current frame and the previous frame, $delt2$ is the gray level difference between the current frame and the next frame, and $delt3$ is the gray level difference between the current frame and the background picture.

3. Experimental results and analysis
The data set used in this article is from "Laboratory for Intelligent & Safe Automobiles". The indoor scenes "Intelligent room" and "Laboratory" and outdoor scenes "Campus" are used to verify the algorithm proposed in this article, and compared with the current commonly used algorithms. The experiment program is written in python3.6 language and runs on a PC with Windows 10 64-bit operating system.

3.1. Feasibility
The data set used in this article comes from "Laboratory for Intelligent & Safe Automobiles". The indoor scene "Intelligent room" and the outdoor scene "Campus" are used to verify the algorithm proposed in this paper, and compared with the current commonly used algorithms. The experiment program is written in python3.6 language and runs on a PC with Windows 10 64-bit operating system.

Among them, the sub-picture (a) in Figure 1-2 is the 100th frame in the video sequence "Intelligent room" and the 63th frame in the video sequence "Campus". The sub-pictures (b), (c) and (d) are the running results of the frame difference method, the running results of the Gaussian mixture model algorithm and the running results of the algorithm, respectively.

3.2. Accuracy
This paper uses two indicators to evaluate and compare the performance of the algorithm, namely the recognition rate (TPR) and the false detection rate (FPR). The calculation method is formula (12) and formula (13).

$$TPR = \frac{TP}{TP + FN}$$  \hspace{1cm} (12)

$$FPR = \frac{FP}{TN + FP}$$  \hspace{1cm} (13)
Among them, TP represents the number of detected pixels belonging to moving objects; FN represents the number of pixels belonging to moving objects that are not detected; TN represents the actual number of pixels belonging to moving objects; FP represents the total number of pixels that do not belong to the background in all pixels of the detected moving target.

Table 1. Recognition rate And False detection rate

| Video sequence of different scenes | Adjacent frame difference method TPR | GMM algorithm TPR | This paper algorithm TPR | Adjacent frame difference method FPR | GMM algorithm FPR | This paper algorithm FPR |
|-----------------------------------|-------------------------------------|-------------------|-------------------------|-------------------------------------|-------------------|-------------------------|
| Intelligent room                  | 0.680                               | 0.790             | 0.890                   | 0.101                               | 0.049             | 0.021                   |
| Campus                            | 0.680                               | 0.765             | 0.870                   | 0.099                               | 0.031             | 0.029                   |

From Table 1, we can conclude that the algorithm proposed in this paper has a higher recognition rate and a lower false detection rate than other algorithms in the table 1. This is mainly because the algorithm in this paper uses dynamic thresholds to filter moving targets, which can adapt to complex environments or changing lighting environments, and overcome the interference of complex environmental noises on moving target detection algorithms.

4. Conclusion
This paper proposes a moving target detection algorithm based on dynamic threshold. By taking the gray difference gradients of three adjacent frames of the video sequence into the foreground moving target segmentation threshold calculation, the dynamic threshold is realized, and the dynamic threshold is used to realize the multi-target motion detection under the complex background or the scene of changing illumination. This article uses the public data set, including indoor environment and outdoor environment to compare the proposed algorithm with other algorithms. It is verified by experiments that the algorithm proposed in this paper has a lower false detection rate and higher accuracy than other algorithms, and can better realize moving target detection in a real environment.

Acknowledgments
Thank you very much Lu Gang for your guidance and revision of this article. This work was supported and funded by the following projects. Shaanxi Normal University Central University Fundamental Scientific Research Business Expenses Innovation Team Project, Project No. GK201801004; Shaanxi Normal University "Double Leaders" Teacher Party Branch Secretary Studio Training Project (No. SDTR201809); Shaanxi Normal University Special Golden Course Construction Project (Computational Thinking) ; Shaanxi Normal University Curriculum Ideological and Political Construction Project (Computational Thinking); Shaanxi Normal University Special Funding for Public Curriculum Reform in 2019, No. 19GGK-JG02.

References
[1] Stauffer C, Grimson W.(1999). Adaptive background mixture models for real-time tracking. In:Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. Fort Collins, Colorado. USA.246~252.
[2] Ding Y. (2014) Detection technology and application of moving target in complex environment. National Defense Industry Press, Beijing.
[3] Xiaofeng, L. , Caidi, X. , Lei, W. , & Lin, T. . (2018). Improved background subtraction method for detecting moving objects based on gmm. IEEJ Transactions on Electrical and Electronic Engineering, 13.
[4] Qu J, XIN Y. (2014) Moving target detection method combining continuous inter-frame difference and background difference. J. Acta Photonica Sinica, 43(07):219-226.
[5] Wei H, Li J, Wu X. (2017) A moving target detection algorithm combining frame difference and ViBe. J. Computer Application Research, 34(05): 1565-1567+1571.

[6] Ju, J., & Xing, J. (2019). Moving object detection based on smoothing three frame difference method fused with rpca. Multimedia Tools and Applications, 78(21), 29937-29951.

[7] Zhang C, Zheng J, Zhang Y, et al. (2017) Moving object detection algorithm based on pixel spatial sample difference consensus. J. Multimedia Tools & Applications, 76(21):22077-22093.

[8] Zheng, F., Shao, L., & Han, J.. (2018). Robust and long-term object tracking with an application to vehicles. IEEE Transactions on Intelligent Transportation Systems, PP(10), 1-13.

[9] Rabah, M., Rohan, A., Mohamed, S. A. S., & Kim, S. H. (2019). Autonomous moving target-tracking for a uav quadcopter based on fuzzy-pi. IEEE Access, 7(99), 38407-38419.

[10] Kim, & Choi. (2019). Indirect time-of-flight depth sensor with two-step comparison scheme for depth frame difference detection. Sensors, 19(17), 3674.