PAPER • OPEN ACCESS

Moving Object Detection Algorithm based on Improved Visual Background Extractor

To cite this article: Xuewen Yang and Tangyou Liu 2021 J. Phys.: Conf. Ser. 1732 012078

View the article online for updates and enhancements.

You may also like

- <u>Detection and tracking of safety helmet in</u> <u>factory environment</u> Xiaoming Sun, Kaige Xu, Sen Wang et al.
- <u>Moving Target Detection Algorithm for</u> Forest Fire Smoke Recognition with Improved ViBe
- Chang Lu, Yichao Cao, Xiaobo Lu et al.
- Four Years of Type Ia Supernovae Observed by TESS: Early-time Light-curve Shapes and Constraints on Companion Interaction Models

M. M. Fausnaugh, P. J. Vallely, M. A. Tucker et al.





DISCOVER how sustainability intersects with electrochemistry & solid state science research



This content was downloaded from IP address 3.144.9.141 on 04/05/2024 at 13:39

Moving Object Detection Algorithm based on Improved Visual Background Extractor

Xuewen Yang^{1,*}, Tangyou Liu^{1,*}

¹School of Donghua University, Shanghai, China

*Corresponding author e-mail: 17521329829@163.com

Abstract. Aiming at the problems that the ViBe algorithm causes error detection when the camera jitter, ghosts appear when processing moving targets, and the detected moving targets are incomplete, an improved ViBe algorithm based on motion compensation is proposed. In the background modeling stage, the motion compensation method based on KLT is used to obtain background model to enhance the robustness of the algorithm to the dynamic background. In the foreground detection stage, combined the background model of the current pixel with 8 neighbor pixel background model, by the double discriminant algorithm to eliminate ghost areas caused by detecting the real background points as foreground points, introduce OTSU algorithm to obtain the optimal threshold for foreground detection. In the post processing stage, filter the connected components for segmentation masks and update models to limit the wrong background points scattered in the foreground.Finally, the connected domain is repaired by morphological opening operation. The results show that the algorithm can effectively eliminate the interference caused by camera jitter, and suppress the ghost phenomenon to obtain more accurate foreground images.

1. Introduction

Moving object detection has always been an important research direction in the field of computer vision and digital image processing. The purpose of moving object detection is to extract moving regions from background images in image sequences. But complex scenes bring huge challenges to the technology, such as dynamic background, camera jitter, ghost, shadows and many other interference factors.

Currently, common moving object detection algorithms include frame difference method[1], optical flow method [2] and background modeling method [3]. Existing background modeling algorithms are roughly divided into two categories [4]: 1) Parameterized background models. Such as Gaussian Mixture Model (GMM) [5], which is a background representation method based on the statistical information of pixel samples; 2) Non-parametric background models. The algorithm such as CodeBook [6,7] is a background extraction algorithm for compressed samples. The Gaussian Mixture model(GMM), codebook and ViBe [8] algorithms are currently excellent background modeling methods.

ViBe (Visual Background extractor) is a moving object detection method proposed by Oliveier Barnich and Marc Van Droogenbroeck[9]. ViBe is an algorithm for pixel-level background modeling and foreground detection. The innovation of this algorithm is applies the random mechanism to the establishment and update of the background model, randomly selects the samples of the pixels to be replaced, and randomly selects the neighboring pixels to update. However, the ViBe algorithm also



Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1732 (2021) 012078 doi:10.1088/1742-6596/1732/1/012078

has shortcomings. ViBe algorithm will misinterpret the background as the foreground and introduce ghosts. And slowly moving objects will merge into the background, resulting in incomplete detection targets.

This paper improves the ViBe algorithm from the following three aspects: 1) Proposed a method based on KLT to obtain the background model after motion compensation, and the impact of the dynamic background is eliminated to the greatest extent; 2) The foreground is detected by double discrimination Obtain the optimal threshold with OTSU, speed up the removal of ghosts; 3) Connected components filtering and morphological filtering to improve the integrity of the foreground image. The results show that the improved method can effectively eliminate the negative effects of dynamic background and camera shake on foreground detection.

2. ViBe algorithm description

In the initial processing stage, the first frame of the video is used to establish a background sample set for each pixel. Then compare the value of all pixels in the image except the first frame in the video to extract the target foreground. Finally, update the background model online.

2.1. Initialize the background model

The ViBe algorithm creates a sample set for each pixel, as shown in Equation(1). The background sample set at the pixel (x, y) is $G_{\nu}(x, y)$, and $G_{\nu}(x, y)$ contains *n* background samples pixel value $v_i (i \in \{0, 1, ..., n\})$.

$$G_{\nu}(x, y) = \{v_1, v_2, \dots, v_{n-1}, v_n\}$$
(1)

Since the adjacent pixels have similar characteristics in time and space distribution, the initialization of the background model $G^0(x, y)$ is based on the first frame M(0) of the video sequence. The selection range of sample pixel values v_i is 8 neighborhoods $N_G(x, y)$ at the pixel (x, y). As shown in Equation(2).

$$G^{0}(x, y) = \{ v^{0}(x_{0}, y_{0}) | (x_{0}, y_{0}) \in N_{G}(x, y) \}$$
(2)

2.2. Foreground object detection

The pixel value at the pixel (x, y) in the frame $M(t)(t \ge 1)$ is used as a detection point v(x, y), and whether the current pixel v(x, y) belongs to the background is determined according to the background model. In the color space, define a sphere $S_R(v(x, y))$ with v(x, y) as it center and R as its radius. Then calculate the number of sample points contained in the intersection of the background sample set $G^{t-1}(x, y)$ and the sphere $S_R(v(x, y))$ [10]. As shown in Figure 1, describes the comparison process of a pixel's background sample set v_i and pixel v(x, y).



Figure 1. Schematic diagram of pixel classification with ViBe algorithm.

2.3. Update the background model

The update strategy of the ViBe algorithm contains three parts: 1) No memory update strategy; 2) Expanding background pixels blocked through foreground pixels by random time re-sampling; 3) Ensure the spatial consistency through the mechanism of expanding the background pixel space[11].

3. Improved ViBe algorithm

3.1. Background model based on motion compensation

In order to reduce the negative impact of camera jitter on detection, image matching tracking algorithm based on KLT[12] is added to the background modeling to establish a motion compensation background model to compensate for the offset caused by camera jitter.

3.1.1. Motion analysis. When the background jitter, the background and moving objects of the current frame have a corresponding offset relative to the previous frame. The current frame image and its background model will produce "disconnection" at the corresponding position. The method of motion compensation is used to compensate this offset in the background model of the current frame. The motion analysis diagram is shown in Figure 2.



Figure 2. Schematic diagram Motion analysis.

3.1.2. Establish motion compensation background model. The main idea of motion compensation is to track some feature points between two consecutive frames through the KLT image matching tracking algorithm, and use the feature points to register the image [13]. The flow diagram of motion compensation algorithm is shown in Figure 3.



Figure 3. Flow diagram of motion compensation algorithm.

Firstly, use the ViBe algorithm to initialize the background model. Then use the KLT method to find N sets of feature matching pixel pairs between two consecutive frames, as shown in Equation (3).

$$\{\{(x_1, y_1), (X_1, Y_1)\}, \{(x_2, y_2), (X_2, Y_2)\}, \dots, \{(x_N, y_N), (X_N, Y_N)\}\}$$
(3)

Due to the cumulative error of image matching tracking, mismatched point pairs are often obtained. In order to improve the accuracy of image matching, this paper uses RANSAC (random sample consensus) [14,15] algorithm to filter the results once, and remove some wrong matching point pairs. And calculate the affine transformation matrix T between the adjacent two frames of images according to the matched point pairs obtained after screening, as shown in (4).

$$T = \begin{bmatrix} \cos\theta s & -\sin\theta s & dx \\ \sin\theta s & \cos\theta s & dy \\ 0 & 0 & 1 \end{bmatrix}$$
(4)

In the matrix, *s* represents the scale ratio of image transformation, the value of *s* in this paper is 1. dx and dy refer to the offset in the *x* and *y* directions respectively. *T* is an affine transformation matrix containing only rotation and translation relationships.

SCSET 2020

Journal of Physics: Conference Series

Finally, the affine transformation matrix T is used to spatially transform the initialized background model to obtain the background model after motion compensation, which is used for the final foreground detection.

3.2. Adaptive threshold segmentation foreground

In order to improve the defects of the ViBe algorithm, a ghost suppression module based on the maximum between-class variance method (OTSU) is introduced in the foreground detection stage of the ViBe algorithm. In the ghost suppression module, the OTSU algorithm is used to calculate the optimal segmentation threshold T_{\min} of the current frame to discriminate pixels. If the number of sample points included in the intersection of the background sample set $G^{t-1}(x, y)$ and the sphere $S_{R}(v(x, y))$ is not less than the threshold T_{\min} , it is considered to be similar to the background sample set $G^{t-1}(x, y)$, and the pixel points v(x, y) are considered to be background points, otherwise the pixel points are considered to be the front points. Compare all the sample pixels according to Equation (5).

$$v(x,y) = \begin{cases} 0, & |S_R(v(x,y)) \cap G^{t-1}(x,y)| \ge T \min\\ 255, & |S_R(v(x,y)) \cap G^{t-1}(x,y)| < T \min \end{cases}$$
(5)

3.3. Double discrimination algorithm based on neighborhood pixels

The ViBe algorithm only uses the first frame to initialize the background model, ignoring the possibility that the first frame contains the foreground. Causing the real background point to be mistakenly detected as a front sight, forming ghosts. In order to avoid generating more uncontrollable factors, ghosts must be suppressed in time to eliminate ghost areas as soon as possible.

The traditional ViBe algorithm ignores the similarity of the neighborhood information in the pixel space. When the background changes rapidly and the background is affected by noise, the background model does not accurately describe the background pixels. In this paper, the foreground sample points are double discriminated based on the background sample set of neighbor pixels, taking full account of the characteristics of the spatial similarity of the neighboring pixels. Flow diagram of double discrimination algorithm as shown in Figure 4.



Figure 4. Flow diagram of double discrimination algorithm.

In the foreground detection stage, the Euclidean distance discrimination between the current pixel v(x, y) and its own background sample set $G^{t-1}(x, y)$ is performed to obtain the sample point set $SG_1(x, y)$, which is intersecting with the background sample set. At the same time, a background sample set $G_N^{t-1}(x, y)$ of neighborhood pixel is randomly selected in 8 neighborhoods. And discrimination is performed between the current pixel v(x, y) and $G_N^{t-1}(x, y)$ to obtain the sample point set $SG_2(x, y)$. As shown in Equation (6), compare two sample point sets to filter out wrong sample points and to further verify the correctness of the foreground set.

$$v(x, y) = \begin{cases} \text{background,} & |SG_1(x, y) \cap SG_2(x, y)| \ge T \text{ min} \\ \text{foreground,} & |SG_1(x, y) \cap SG_2(x, y)| < T \text{ min} \end{cases}$$
(6)

SCSET 2020		IOP Publishing
Journal of Physics: Conference Series	1732 (2021) 012078	doi:10.1088/1742-6596/1732/1/012078

3.4. Connected components filtering and morphological processing

In the conservative update strategy, different connected components filtering methods are used to process the segmentation and update masks in this paper. Remove the foreground blobs whose area less than or equal to 10 pixels in the segmentation masks, and fill the foreground holes whose area less than or equal to 20 pixels; fill the foreground hole area with area less than or equal to 50 pixels in the update models, and keep all the foreground for the update masks speckle, because the foreground value in the conservative update strategy of ViBe detection is inserted into the background model. Use morphological operations to overcome holes, eliminate outliers, and improve the integrity of goals.

4. Experiments and analysis

4.1. Result analysis

The results of this paper are all run on a computer with an Intel Core i7-10710U CPU@1.10GHz 1.61GHz frequency and 16 GB of memory. The operating environment is VisualStudio 2015, combined with the open source OpenCV library. Frame Difference(FD), GMM, traditional ViBe and improved ViBe algorithm are used as comparative experiments. This paper selects the video sequences from the data set provided by CDW-2014[16]. The results of the four algorithms are shown in the Figure 5.



(b)traffic(Frame 352, FD, GMM, ViBe, Ours)

Figure 5. Experimental results of 4 algorithms in the camera jitter scenes.

In the badminton scene, camera jitter has a great interference on object detection. The FD, GMM and ViBe algorithm are all affected by jitter, and cannot obtain correct object prospect. The improved algorithm in this paper has a excellent detection effect. The details are also handled very well. In traffic scenes, due to camera shake, the GMM algorithm detects roads as moving objects, and there are more erroneous foreground points. The algorithm in this paper can largely eliminate the negative effects of camera jitter.

In addition to experimenting on the camera jitter video sequences in the data set, also detected airport runway video with camera jitter problem. The detection results of frame 81 and frame124 are shown in Figure 6.



(a)skating(Frame 124, FD, GMM, ViBe, Ours)

Figure 6. Flow diagram of double discrimination based on neighboring pixels.

SCSET	2020

IOP Publishing

The improved ViBe algorithm effectively removes the interference caused by camera jitter. In summary, the improved ViBe algorithm is significantly better than other algorithms in most cases, and the detection effect is also more robust in complex scenes.

4.2. Evaluation and analysis

In order to comprehensively evaluate the algorithms, this paper uses three parameters as evaluation indicators. As shown in Equation(7).

$$\begin{cases} precision = \frac{TP}{TP + FP} \\ recall = \frac{TP}{TP + FN} \\ F - M = \frac{2^* precision^* recall}{precision + recall} \end{cases}$$
(7)

Precision represents the proportion of actual correct front attractions in all foreground judgments, *recall* represents the proportion of actual correct front attractions in the truth map, FM (F-Measure) is the weighted harmonic average of *precision* and *recall*, Comparison of different indicators of the four algorithms as shown in the figure 7.



Figure 7. Comparison of different indicators of the four algorithms.

In summary, through comprehensive evaluation of the above algorithms by F-M parameters, the experimental results of our algorithm are improved by 32.0%, 37.0%, and 4.0% compared with the FD, GMM, and ViBe algorithms, respectively. The algorithm in this paper has the best comprehensive performance in all aspects.

5. Conclusion

In view of the dynamic background caused by camera jitter, proposed an improved ViBe algorithm based on the motion compensation background model. Firstly, the background model is established through the motion compensation algorithm, and introduced the adaptive threshold in the foreground detection stage. The double discrimination method suppresses the false background points scattered in the foreground. Through the filtering operation and morphological processing of the connection components, the foreground hole areas are filled, and solitary points are filtered out. The experimental results on the airport runway video and camera jitter videos collected by the data set show that our algorithm has good comprehensive performance, can effectively suppress ghosts and remove jitter interference, and obtain more accurate foreground images. However, the algorithm in this paper still has defects in the detection of shadowed objects, and the robustness in this aspect needs to be improved.

References

- Z. Yaoming, W. Chengdong, Z. Yunzhou and F. Sheng, Realization of moving object detection and tracking algorithm based on frame difference method and particle filter algorithm, 2017 29th Chinese Control And Decision Conference (CCDC), Chongqing, 2017, pp. 161-166, doi: 10.1109/CCDC.2017.7978085.
- [2] Fernando W S, Udawatta L, Pathirana P N, et al. Identification of moving obstacles with Pyramidal Lucas Kanade optical flow and k means clustering[C]. international conference on information and automation, 2007: 111-117.
- [3] Benezeth Y, Jodoin P, Emile B, et al. Review and evaluation of commomly-implemented background subtraction algorithms[C]//19th International Conference on Pattern Recognition , 2008:1-4.
- [4] Zhuang Zhemin, Zhang Congyou, Yang JInyao, et al. Investigation on visual background extractor based on gray feature and adaptive threshold[J]. Journal of Electronics & Information Technology, 2015,37(2):347-352.
- [5] Chen Yin, Ren Kan, Gu Guohua, et al. Moving object detection based on improved single Gaussian background model[J]. Chinese J Lasers,2014,41(11):1109002.
- [6] Kim K, Chalidabhongse T, Harwood D, et al. Background modeling and subtraction by codebook construction[C]//2014 IEEE International Conference, 2004, 5:3061-3064.
- [7] Kim K, Chalidabhongse T, Harwood D, et al. Real-time foreground-background segmentation using codebook mode[J]. Special Issue on Video Object Processing, 2005,11(3):172-185.
- [8] Barnich O, Van Droogenbroeck M. ViBE: A powerful random technique to estimate the background in video sequences[C]. international conference on acoustics, speech, and signal processing, 2009: 945-948.
- [9] Barnich O, Van Droogenbroeck M. ViBe: A Universal Background Subtraction Algorithm for Video Sequences[J]. IEEE Transactions on Image Processing, 2011, 20(6): 1709-1724.
- [10] Qu Z, Yi W, Zhou R, et al. Scale Self-Adaption Tracking Method of Defog-PSA-Kcf Defogging and Dimensionality Reduction of Foreign Matter Intrusion Along Railway Lines[J]. IEEE Access, 2019: 126720-126733.
- [11] Gao, Jun & Zhu, Honghui. Moving object detection for video surveillance based on improved ViBe[C]// CCDC. 2016:6259-6263.
- [12] Ramadhan J. Mstafa and Khaled M. Elleithy. A video steganography algorithm based on Kanade-Lucas-Tomasi tracking algorithm and error correcting codes[J]. Multimedia Tools Appl. 2016, 75, 17, 10311 - 10333.
- [13] Zheng H., Chen C. Feature Point Detection and Target Tracking Based on SIFT and KLT[C]//Communications in Computer and Information Science. Springer, Singapore . 2018, vol 849.
- [14] Schnabel R, Wahl R, Klein R, et al. Efficient RANSAC for Point-Cloud Shape Detection[J]. Computer Graphics Forum, 2007, 26(2): 214-226.
- [15] Laraalvarez C, Romero L, Flores J F, et al. A Simple Sample Consensus Algorithm to Find Multiple Models[C]// Iberoamerican congress on pattern recognition. 2009: 918-925.
- [16] Wang Y, Jodoin P, Porikli F, et al. CDnet 2014: An Expanded Change Detection Benchmark Dataset[C]// computer vision and pattern recognition. 2014: 393-400.