Moving object detection of assembly components based on improved background subtraction algorithm

K Zhang¹, S Tong¹*, H Shi¹, G Yue¹ and J Zhao¹
¹College of Mechanical Engineering, Shenyang Jianzhu University, China.

*Email: tongshenghao@sjzu.edu.cn

Abstract. In the automatic assembly of mechanical components, it is difficult to guarantee the quality of assembly image acquisition caused by the change of assembly environment. Firstly, the color space is compared and analyzed, and histogram equalization is used to preprocess the image to reduce the impact of image brightness on subsequent processing. Then, based on the texture features of the image, an image reconstruction idea from coarse to fine is created, and an image representation method from macro to micro is established to suppress the pseudo texture and noise texture in image processing. On this basis, an adaptive threshold algorithm based on the target distance is proposed to improve the detection accuracy of the assembly target. Experimental results show that the proposed algorithm can significantly improve the detection accuracy of moving objects in assembly video frames.

1. Introduction
Moving object detection has been widely used in automatic positioning, security monitoring and other scenes. It is particularly important for the development of computer vision to correctly detect moving objects from video [1, 2]. Background modelling is the core part of fixed camera video monitoring, which is used to segment moving objects from the background [3]. Therefore, background modelling in video surveillance is also called moving object detection. In this technology, the moving object is called foreground object, because its pixels are not suitable for background model, and can be extracted by "background subtraction" [4, 5].

In recent years, scholars have been studying the moving object detection algorithm using image features and color parameters for feature extraction [6]. In order to segment the moving foreground in real-time and accurately in complex scenes such as illumination change and dynamic background interference, Jin [7] proposed a moving object detection method based on Region Histogram modelling in color attribute space, aiming at the shortcomings of traditional methods based on color features and pixels. Firstly, the RGB color space is mapped to a more robust low-dimensional color attribute space, and the histogram is established in the local range of the pixel with the color attribute as the feature. At the same time, the spatial information of the pixel in each partition of the histogram is recorded, and the K spatial histograms are used to form the background model of each pixel, and each histogram is giving different weights according to its matching degree. However, it is difficult to set the weight parameters in the actual operation of this method, and the results are greatly influenced by the weight parameters. Yin [8] proposes an anomaly detection algorithm based on the background of regular changes. Firstly, the region segmentation algorithm based on three frame difference method is used to segment the image into static background region and dynamic background region including moving objects. At the same time of segmenting static background area and dynamic background area,
Surendra algorithm based on three frame difference is used to extract static background area information and establish background model. Ding [9] proposed vibe background modelling with color and edge features to solve the problem of single feature in traditional background modelling methods. The problem of noise, breakpoint and internal hole in moving object detection results in three frame difference method was solved, and the results of image processing were compensated based on morphological processing method. But he did not solve the problem of image noise caused by the movement of the target in the video image. Although traditional background subtraction algorithms have been widely used in different environments, they still face various challenges [10]. One of them is that the foreground object is often incomplete with a lot of noise, which means that it contains holes or scattered fragments [11, 12]. These defects can cause false detection and affect the efficiency of advanced applications such as target location and appearance recognition [13, 14, 15]. To solve this problem, a more stable background modeling method is needed to extract the complete and clear foreground objects.

In this paper, from the perspective of improving the image threshold, the research of moving object detection in component assembly location video image is carried out. In the assembly location video of components, different image frames will produce different sizes and positions of targets, and the size and position parameters of targets will have different effects on the extraction of target features. Therefore, the corner points are set on the edge of the target feature extracted from the previous frame image, and the distance norm model of the pixels and corner points in different regions of the image is established. Based on this, an adaptive threshold algorithm based on the target distance is proposed to suppress the noise and improve the detection accuracy of the moving target.

2. Fundamental principles
Image threshold processing is an essential step in many visual tasks. The purpose of threshold processing is to classify pixels as "dark" or "bright". Image thresholding the segmentation of a digital image based on specific features of the pixel, such as intensity values. To create a binary image, categorizing each pixel into one of two categories. Integral image is a tool. When there is a mapping function from pixel to value and calculate the sum of function values in a rectangular area, integral image is a very efficient tool. In the case of no integral image, it need to calculate the sum of the mapping function of each pixel in each rectangular window by calculating the function value of each pixel separately and finally stacking. When computing the sum of multiple overlapping rectangular windows, the use of integral images can significantly reduce the amount of computation.

In order to calculate the integral image, it should store the sum of all the items of the left and upper pixels in each position. For each pixel, calculate integration image using the following equation [4].

\[ I(x, y) = f(x, y) + I(x - 1, y) + I(x, y - 1) - I(x - 1, y - 1) \] (1)

The sum of the functions of any rectangle with upper left and lower right corners would be calculated as [6].

\[ \sum_{x=1}^{x_2} \sum_{y=1}^{y_2} f(x,y) = I(x_2, y_2) - I(x_2, y_1 - 1) - I(x_1 - 1, y_2) + I(x_1 - 1, y_1 - 1) \] (2)

The job of this method is to compare with nearby pixels. In addition, because the neighborhood samples are not evenly distributed in all directions, it is not a good representation to use the average step size in the whole image. This kind of problem is solved by using integral image, and the same output is produced for different processing methods. This is not to calculate the average value of the last s pixels, but to calculate the average value of the s×s pixel window centered on each pixel. This is a better average because it takes neighboring pixels on all sides into account. The time of average calculation using integral image is linear. The integral image is calculated when the image is traversed for the first time. In the second pass, use the integral image to calculate the s×s average value of each
pixel, and using the time as a constant, then compare. Black or white will be set after the pixel value is compared.

3. Algorithm for moving object detection

3.1. Feature extraction based on background subtraction

For a perfect assembly location environment, the gray value of each pixel in the image is in line with the random probability distribution when there is no redundant moving target, the illumination does not change and other environmental factors do not change. Because of the actual engineering environment and in the process of image acquisition, it is inevitable to introduce noise artificially. These gray values can be kept in a certain gray value, and random excitation within a certain range. This kind of scene is called "background". Setting up background model and updating background are the most important problems in background subtraction.

Create a background image frame $B$, note that the current image frame is $A$, and the gray values at the pixel points of the current image frame and the background image frame are respectively recorded as $B(i, j)$ and $A(i, j)$. According to equation (1), the difference image $C$ is shown as:

$$C_n(i, j) = |A(i, j) - B(i, j)| \quad (3)$$

After the background model is established successfully, the next step is target detection. Set threshold $\sigma$, according to equation (2), binary image is obtained by binary processing of the difference image pixels $T_n$. And, the image $T$ with complete moving target can be obtained.

$$T_n(i, j) = \begin{cases} 1, & C_n(i, j) > \sigma \\ 0, & \text{else} \end{cases} \quad (4)$$

The system needs to update the background model in real time when the background is slightly changed due to light changes. In order to quickly adapt to changes in the scene, the following formula is carried out.

$$B_{n+1} = B_n - \frac{B_n}{n+1} + \frac{\alpha_{n+1}}{n+1} \quad (5)$$

The operation process is shown in equation (3). ‘$\alpha$’ represents the update weight coefficient. The general rule of background update is that if the current point is determined as a background point, the background value of the point is updated, otherwise it is unchanged. This can save the update time and speed.

3.2. Adaptive threshold modeling

High threshold will lead to fragmentation of foreground objects, while low threshold will lead to noise of pseudo foreground objects. To solve this problem, we propose an adaptive threshold algorithm based on target distance to improve the accuracy of target feature extraction. Different image frames will produce different sizes and positions of targets, and these parameters will have different effects on the extraction of target features. As shown in figure 1, we first divide the image into different blocks, the corner points on the edge of the target feature of the previous frame must be set, the distance norm model between pixels and corner points of different blocks of the image must be established, and the adaptive threshold algorithm must be established to complete the target extraction of the image.
In order to ensure real-time, the distance norm between each pixel in the image and the corner of the target edge of the previous frame cannot be calculated. Therefore, use the method of dividing the image into different blocks to calculate the distance norm between the center pixel of each block and the edge corner of the previous image, which can not only improve the accuracy of feature extraction, but also reduce the amount of calculation. Here, it use P to represent the set of image block pixel center points, and Q to represent the set of target edge corner points. Distance model \( D(P, Q) \) is defined as follows:

\[
D(P, Q) = d(P, Q) + d(P, Q) = \min(q \in Q, p \in P) \|p - q\| + \max(q \in Q, p \in P) \|q - p\|.
\]

\( \|\cdot\| \) represents the distance norm between two sets of points. On this basis, we define the adaptive threshold as follows:

\[
K_p = \frac{\alpha_{n+1}}{D(P, Q_n)} \tilde{K}.
\]

\( \tilde{K} \) represents setting a threshold. Then the feature extraction can be expressed as follows:

\[
b_{i,j} = \begin{cases} 
0, & c_{i,j} < K_p \\
1, & \text{or} \end{cases}
\]

Among them, \( \alpha \) represents the updated weight coefficient, \( b_{i,j} \) represents the extracted image feature value, and \( c_{i,j} \) represents the differential image feature value.

4. Experiment
In order to verify the performance of the proposed algorithm, an assembly component positioning experimental platform is built in the laboratory, as shown in figure 2. On the acquired assembly component location video data set, the algorithm proposed in this paper is compared with the traditional background subtraction algorithm, literature [7] and literature [9], and the results are shown in figure 3.
Figure 2. Assembly platform.

Figure 3. Comparison results: (a) input image; (b) ours; (c) literature [7]; (d) literature [9]; (e) traditional.

In column (a), the red rectangle box indicates the target to be extracted. From column (b), it can be seen that the target edge extracted by this method is clear, the internal defects are small, the external noise is small, and the detection results are obviously better than other detection methods. As column (c), it shows that there is external noise near the target extracted by literature [7], and the extraction accuracy is lower than that of this method. As column (d), it shows that the internal defects and external noise of the target extracted by the algorithm in literature [9] are more than other methods, and the extraction accuracy is lower. From the detection results of (e) column, it can be seen that the
target extracted by traditional background subtraction algorithm has more external noise, and there will be incomplete extraction of the target, which shows that the method has poor results in external interference, and the accuracy of target feature extraction is low.

5. Conclusion
In this paper, we aiming at the problem of low target detection accuracy of background subtraction algorithm based on fixed threshold, the assembly location video image has been studied. This paper introduces the principle of background subtraction algorithm, carries out background modeling for assembly target, sets corner points on the edge of target feature extracted from the previous frame image, establishes the distance norm model of pixels and corner points in different regions of this frame image, and proposes an adaptive threshold algorithm. In the laboratory, the assembly positioning experiment platform is built, the video image data set is made, and the assembly positioning images at different times are selected for comparative experiments. The results show that the algorithm in this paper greatly improves the target detection performance of the assembly video image, and has a strong ability to suppress image noise.

Acknowledgments
This work was supported by the National Key R&D Plan of China, grant number 2017YFC0703903, 2017YFC0704003 and the Shenyang Science and Technology Plan Project (Y19-1-004).

References
[1] He Y 2016 Moving Object Detection Based on Improved Frame Difference Method and Background Difference Method Yanshan University
[2] Liu T 2016 Research on Moving Object Detection Algorithm Based on Background Modelling Zhengzhou University
[3] Liu Z, He S and Hu W 2017 Moving object detection of video sequence based on background subtraction Computer Application 37 (6) 1777-81
[4] Hu Z, Xi Y and Tong J 2018 Adaptive moving target detection in the background of uneven clutter Command Information System and Technology 9 28-32
[5] Zhang Y and Liu Y 2017 Moving object detection based on frame difference method and background difference method Computer Technology and Development 27 (2) 25-28
[6] Li Y, Xu X and Chen J 2017 Application of background update of dynamic feature block matching in motion detection Journal of Instrumentation 38 (2) 445-53
[7] Jin J, Dang J and Wang Y 2019 Spatial histogram background modeling of regional color attributes Chinese Journal of Image Graphics 714-23
[8] Yin R, Yang J and Lu W 2019 Research on anomaly detection algorithm based on the background of regular change Control Engineering 26 (8)
[9] Ding Z and Lu W 2019 Moving object detection algorithm based on vibe background modeling. Computer System Application 28 (04) 187-91
[10] Di C 2017 Research on Video Detection and Tracking Methods of Moving Objects Based on Features University of Chinese Academy of Sciences
[11] Yin H, Chen B, Chai Y et al. 2016 Overview of vision based object detection and tracking Journal of Automation 42 (10) 1466-89
[12] Wang D, Fan Y and Wang X 2017 A motion detection algorithm based on gradient and vibe algorithm Television Technology 41 (3) 100-04
[13] Zhan Z, An Y and Cui W 2017 Image threshold segmentation algorithm and comparative study Information Communication (4) 86-89
[14] Chen J 2017 Research on Threshold Segmentation of Non-uniform Illumination Image Hunan University of Technology
[15] Hu Z, Fan W and Yu G 2017 Adaptive threshold image foreground extraction based on OpenCV Photoelectric Technology Application 32 (1) 56-59