A Target Detection Method Based on Statistical Mean Value Difference Model

Fei Lei\textsuperscript{1, a}, Kai Long\textsuperscript{1, b} and Xueli Wang\textsuperscript{1}

\textsuperscript{1}Beijing University of Technology, Beijing, China

\textsuperscript{a}myislj@126.com; \textsuperscript{b}longkai225@163.com

Abstract: In the field of oil pipeline monitoring applications, the guarantee of high accuracy target detection is required. Aiming at the problem of moving target detection under static camera, a method of mean background model based on probability and statistics is proposed. The background model is quickly constructed. The background threshold is adaptively used to eliminate the influence of illumination changes. The target is detected quickly and accurately. Finally, the proposed target is proposed. The background update method updates the background model in real time. The simulation results show that the better detection results can be obtained under the influence of light changes. The background model has strong stability and adaptability, and can be applied to the real-time monitoring system of the field oil pipeline.

1. Introduction

Due to its safety, economy and speed, the oil pipeline has received attention and rapid development in China in recent years. While the oil pipeline is developing rapidly, some lawless elements have made huge profits by stealing national crude oil, which has hindered the development of the oil pipeline transportation industry and the loss of state property [1]. Therefore, it is necessary to monitor multiple parts of the oil pipeline by means of computer vision image processing technology. Common moving target detection methods are optical flow method [2], adjacent frame difference method [3], and background difference method [4]. Although the optical flow method can detect moving objects independently, it is very time consuming and difficult to meet real-time. The adjacent frame difference method can adapt to the dynamic environment quickly, but the segmented moving objects are incomplete. The background difference method can make up for the first two. The shortcomings of the methods are the best in terms of speed and effect.

In this paper, an improved algorithm for mean background modeling is proposed. The mean model background is constructed by mathematical modeling method of probability and statistics. The threshold is adaptively selected according to the illumination change of the image. The image to be detected is compared with the background image, and then the foreground target is extracted. In the process of constructing the background model, the background image is updated in real time, effectively eliminating the interference of environmental factors, and thus obtaining an ideal background image.

2. Background model

2.1. Mean background model

The mean background model is the averaging of multiple images based on time, that is, the
background image is obtained by obtaining the average of the sum of the continuous video images over a period of time.

Initialize the image sequence, use the nearest neighbor interpolation method to adjust the image size to $H \times W$, and the mean background model formula is as follows:

$$ B_k(i, j) = \frac{1}{N} \sum_{i=1}^{N} A_k(i, j) $$

In formula (1): $B_k(i, j)$ is the pixel value of the background image frame at time $k$ at $(i, j)$ position; $A_k(i, j)$ is the pixel value of the input image frame at time $(i, j)$ at time $k$; $N$ is the total number of frames of continuous images.

2.2. Improved mean background model

The mean background model is sensitive to ambient light changes and branching, and sometimes “smear”. In order to reduce the influence of these factors, a probabilistic method is used, and each of the pixel positions may be generated by using a statistical histogram. The number of occurrences of the pixel value is counted, and several gray values with a large number of occurrences are added and averaged as the pixel value of the background.

Assuming a sequence of $N$ frames of images, the pixel value that may occur at each pixel location is $c = \{c_1, c_2, ..., c_l\}$. The number of occurrences of each pixel value is $m = \{m_{c_1}, m_{c_2}, ..., m_{c_l}\}$. The total number of times is equal to $N$, namely:

$$ m_{c_1} + m_{c_2} + ... + m_{c_l} = N $$

(2)

Judging each pixel value traversal, if the number of occurrences meets the requirements, accumulate the pixel value, the formula is as follows:

$$ m' = m' + m_c \cdot \frac{m_c}{N} > \varepsilon \quad \text{(Where } c = \{c_1, c_2, ..., c_l\}, m' \text{ initial value is 0} ) $$

(3)

In equation (3), $\varepsilon$ is the empirical value, generally about 0.3~0.4. The improved mean background model is as follows:

$$ B_k(i, j) = \frac{B(i, j)}{m'} $$

(4)

The image to be detected is subtracted by the background model of the structure, and the pixel value after subtraction is binarized. It is important to set the segmentation threshold. If the threshold is too large, some foreground pixels will be judged as the background; if the threshold is taken The value is too small, and some background pixels will be judged as foreground. In this paper, the adaptive threshold selection method is adopted to make the threshold $T$ with the largest variance between classes the most suitable threshold. The pixels smaller than $T$ are set to 0, and the pixels larger than $T$ are set to 255 to obtain the foreground image. The formula is as follows:

$$ D(i, j) = \begin{cases} 255, & |I_k(i, j) - B_k(i, j)| > T \\ 0, & |I_k(i, j) - B_k(i, j)| \leq T \end{cases} $$

(5)

3. Adaptive background differential target detection

3.1. Process of target detection algorithm
The basic steps of target detection: image frame analysis of the input video, the background model is established by the mean method, the foreground area is obtained by adaptive threshold background difference, and the post-processing is finally performed and the detection target is identified. The flowchart is shown in Figure 1.

3.2. Specific method of detection

- Obtain an input image as a video sequence image \( \{A_k(i,j)\}_{k=1}^N \) at each of the first N days of the current time, where \( k \) is the sequence of frames and \( N \) is the total number of sequence images.
- The statistical histogram is used to count the number of pixel values that may occur at each pixel position, and several pixel values with a large number of times are summed and averaged as the background pixel of the pixel position, that is, the background model \( B_k(i,j) = \frac{B(i,j)}{m} \) is obtained.
- Each pixel of the image is traversed, and each pixel is binarized. A pixel larger than the threshold \( T \) is set as a foreground; a pixel smaller than the threshold \( T \) is set as a background, that is, a binary foreground image is obtained.
- Real-time update of the built background model, background update rate is \( \alpha \in (0,1) \). The background update model is \( B_{k+1}(i,j) = \alpha B_k(i,j) + (1-\alpha)I(i,j) \). The next frame background image is related to the current frame image and the previous image.
- Identify the detection target of the foreground image and determine the type of the target.

3.3. Selection of threshold \( T \)

The value of the image binarization segmentation threshold is very important, and the use of a fixed threshold cannot adapt to the effects of ambient illumination changes. If the split value is too small, the background pixel will be mistaken for the foreground pixel; if the value is too large, some foreground pixels will be mistaken for the background pixel.

The largest inter-class variance method was first proposed by Otsu, also known as the Otsu threshold algorithm [5]. The idea of the algorithm is to use the principle of least squares to determine an optimal threshold, so that the image is binarized, and the variance between the foreground and background pixels is guaranteed to be the largest.

Assuming that the total number of pixels of the image to be processed whose gray level is \( L \) is \( N \), and the total number of pixels whose gray level is \( i \) is \( n_i \), the frequency at which each gray value appears is expressed as:

\[
P_i = \frac{n_i}{N}, i = 1, 2, ..., L-1
\]
It is assumed that the image is divided into foreground and background regions with grayscale K as the threshold, pixels with gradation of $0 \sim k$ and pixels with gradation of $k+1 \sim L-1$ constitute region B (background) and F (foreground), the frequencies are:

$$P_B = \sum_{i=0}^{k} P_i \quad \text{and} \quad P_F = \sum_{i=k+1}^{L-1} P_i$$

Then the average gray levels of area B and area F are:

$$\mu_B = \frac{\sum_{i=0}^{k} i \ast P_i}{P_B} \quad \text{and} \quad \mu_F = \frac{\sum_{i=k+1}^{L-1} i \ast P_i}{P_F}$$

Then the inter-class variance of the image can be defined as:

$$\mu_k = P_B \left( \mu_B - k \right)^2 + P_F \left( \mu_F - k \right)^2$$

The traversal method is used to obtain a threshold T that maximizes the variance between classes, which is the most appropriate threshold.

3.4. Background update

The background model is constructed based on a sequence of images over time. Over time, it may be disturbed by factors such as changes in illumination, branching, and sudden motion of stationary scenes. In order to make the modified background model have certain adaptability to external changes, it is necessary to update the background model to overcome the interference of these factors.

The basic method of background update is to traverse each pixel of the image and determine that the pixel is the foreground pixel, and keep the original background pixel value without updating; if it is judged as the background pixel, it is updated, and the background update formula [6] is:

$$B_{k+1}(i, j) = \alpha B_k(i, j) + (1 - \alpha)I(i, j)$$

In equation (10), $\alpha$ is the background update rate, and its range is (0,1), where $I(i, j)$ is the current frame image, $B_k(i, j)$ is the gray value of the current background frame image, and $B_{k+1}(i, j)$ is the gray value of the background frame image of the next time. It can be seen from equation (10) that the background model is related to the current frame image and the previous frame image, effectively avoiding sudden phenomena, and will remain relatively stable over time, and can adapt to the effects of illumination changes, branch swings and the like.

4. Experimental results and analysis

4.1. Evaluation standard

The evaluation criteria of image quality are mainly divided into subjective and objective evaluation. The subjective evaluation is mainly based on the human eye to observe the image, and it is evaluated by the feeling, which is easily affected by personal factors, so the objective evaluation standard is generally used.

The objective evaluation standard is mainly to establish a mathematical model for the image, and measure the information index through the specific formula. This paper studies the information entropy [7] and the edge ratio of the whole image.

Information entropy is the probability that the random variable appears. The size of the information is measured by the magnitude of the probability. Generally, the amount of information can only be represented by the change. Definition of information entropy: Assume that the gray level of the image
is \( N \), and the value of the random variable is \( p_{si} = p\{X = x_i\}, i = 1, 2, ..., N \). \( x_i \) represents the gray value that appears, and the information entropy is expressed as:

\[
H[S] = -\sum_{s=0}^{N-1} p_s \log p_s
\]  

The edge ratio represents the edge information of the image. When the number of targets is large, the edge ratio of the image is larger, and the more edge information is retained. The edge ratio is expressed as:

\[
R = \frac{P_{\text{edge}}}{(W \times H)}
\]

In equation (12), \( P_{\text{edge}} \) represents the number of edge information with image size \( W \times H \). In this paper, the sobel edge detection algorithm [8] is used to detect the number of edge pixels in the image.

### 4.2. Experimental results and analysis

In order to verify the effectiveness of the proposed algorithm, simulation experiments were carried out on the experimental data collected. The experimental platform runs on the operating system Windows7 Professional, the video image resolution is 320×280, with the suffix name in avi format, and the programming environment is vs2010 and the opencv2.4.10 machine vision library.

The improved mean background modeling effect is shown in the following figure:

![Figure 2](image)

As shown in the above figure, Figure 2(b) and (e) are background models constructed by accumulating the mean values of a sequence of pixels. It can be seen that the branches are swaying, the wires are swaying, and the motion of the moving objects will appear "smear", which will reduce the background image quality. Figure 2(c) and (f) show an improved mean background image, which significantly suppresses the effects of these factors and achieves better results.

As can be seen from Figure 3 below, the effect of using the adaptive threshold is better than selecting a fixed threshold for binary processing, and the adaptive threshold can be selected to select the most appropriate threshold for segmentation. In Figure 3(c), some interfering pixels appear in the image, and some background pixels are mistaken for foreground, while graph Figure 3(d) achieves better experimental results. It can be seen from Table 1 that the improved algorithm adopted in this paper is slower in processing time, but the richness of the internal state of the constructed background image and the edge information achieve better results.
Figure 3. Target Detection. (a) Original image, (b) Grayscale frequency histogram, (c) Fixed threshold detection, and (d) OTSU algorithm detection.

Table 1. Experimental comparison of background modeling algorithms

| Detection image | Background modeling method | Algorithm processing time (ms/f) | Information entropy | Edge ratio |
|-----------------|----------------------------|----------------------------------|---------------------|-----------|
| **Figure 2.** Scene one the 200th frame image | Mean background modeling | 1.120 | 7.0140 | 0.0691 |
| | Background modeling of statistical mean | 1.323 | 5.5776 | 0.0731 |
| **Figure 2.** Scene two the 256th frame image | Mean background modeling | 0.974 | 6.2420 | 0.0715 |
| | Background modeling of statistical mean | 1.212 | 7.6661 | 0.0761 |

5. Conclusion
The background model based on statistical mean is improved over the mean background model method. The background model is simpler to implement, the edge information of the image is better preserved, and the negative effects caused by sudden changes in background can be better eliminated. The model is more stable and improves the accuracy of target detection.

References
[1] Xufeng Yu. Research on Pattern Recognition Technology of Burglary Police in Anti-theft Detection of Oil Pipeline [D]. Nanjing Aerospace University, 2004, 1.1.
[2] Hiraia A, Fuse K, Komatsu N, et al. Accurate Estimation of Optical Flow for Fully Automated Tracking of Moving-Objects within Video Streams [J]. IEEE International Symposium on Circuits and Systems, 1999, 4: 515-519.
[3] Lipton A, Fujiyoshi H, Patil R. Moving Target Classification and Tracking from Real-time Video [C] / Applications of Computer Vision, 1998, WACV'98. Proceedings, Fourth IEEE Workshop on Princeton, NJ, USA: IEEE, 1998: 8-14.
[4] HAN Hong-zhe. Adaptive Background Modeling with Shadow Suppression [J]. The IEEE 6th International Conference On Intelligent Transportation Systems, 2003, 1: 720-724.
[5] Qian Wang, Zhong Peng, Li Liu. An Image Segmentation Algorithm Based on Adaptive Threshold [R]. Transactions of Beijing Institute of Technology, 2003, 23(4): 521-524.
[6] Chong Wang, Zhihong Xi, Chunli Xiao. Moving Target Detection Method Based on Background Difference [J]. Applied Science and Technology, 2009, 36(10): 16-19.
[7] Junbo Wang, Yinglei Song, Bing Zhang, Taoying Sun. Target Detection for Background Model Selection [J]. Modern Electronics Technique, 2018, 41(13): 34-37.
[8] Meijing Zhang. Research on sub pixel edge detection technology [D]. Shenyang: Shenyang Ligong University, 2013.