An Automatic Moving Object Detection Algorithm for Video Surveillance Applications

Xiaoshi Zheng, Yanling Zhao, Na li, Huimin Wu
Shandong Computer Science Center
Jinan, China
zhengxs@keylab.net

Abstract—Moving object detection is a very important step in video surveillance. And frame difference algorithms are suitable for these applications. First of all, an automatic threshold calculation method was proposed according to statistic information to obtain moving pixels of video frames. Then moving regions can be formed by morphological operations. At last, the nearest distance of two regions was proposed and it was satisfying for region combination. The proposed algorithm is automatic and efficient in intelligent surveillance applications.

Keywords —moving object; threshold; frame difference; region combination; nearest distance

I. INTRODUCTION

Moving objects detection in video streams is the first relevant step of information extraction in many computer vision applications, including traffic monitoring, automated remote video surveillance, people tracking, etc. And it is an important basis of later steps, such as recognition, classification, and activity analysis [1]. Although a lot of studies have been proposed in recent years, the subject is still challenging [2].

Conventional approaches to moving objects detection include frame difference (or temporal difference) algorithms, background subtraction algorithms, optical flow algorithms, and statistical learning algorithms. Optical flow and statistical learning algorithms have much computational complexity and are not suitable for video surveillance applications. While background subtraction is very sensitive to light changes. In contrast, frame difference algorithms discern static objects (having null differences) from moving objects (having no-null differences), and they are simple and easy to be implemented. So they are considered to be suitable for video surveillance applications.

This paper presents an automatic moving objects detection algorithm based on frame difference and region combination. Moving regions were obtained automatically by frame difference with an adaptive threshold. Then region combination was done according to the nearest distance. Experiments results show the proposed algorithm is automatic and efficient in moving objects detection for video surveillance application.

II. PREPARATIONS

A. Median Filtering and Select Color Component of Frames

Before processing the frames of video, we perform 2-D median filtering first to smooth frames and filter noise. Though many algorithms are based on gray images, they are not better than those which only use one component of images (such as R, G and B component). In this article, we choose R component of each frame to detect moving objects in order to shorten runtime.

B. Threshold Selection of Binary Process

Selection of threshold is a crucial step for binary process. The value of threshold is changeable with frames of video, and it will affect the efficiency of the whole proposed algorithm if it is too large or small. In order to obtain an adaptive threshold automatically from frames, many people extract statistic characters (such as mean, standard deviation, etc) [3, 4].

In this paper, we propose an automatic threshold selection method according to calculation described in [4]. Threshold value can be calculated as following:

\[ u = \frac{1}{r \times c} \sum_{i=1}^{r} \sum_{j=1}^{c} FD(i, j). \]

\[ \sigma = \sqrt{\frac{1}{r \times c} \sum_{i=1}^{r} \sum_{j=1}^{c} \left( FD(i, j) - u \right)^2}, \] (1)

\[ T = 1.5 \times \sigma \]

In which, \( FD \ (i, j) \) is the difference of pixel \( (i, j) \) between two successive frames, and \( r, c \) is the size of frames.

III. PROPOSED ALGORITHM

A. Moving Pixels

Frame difference is obtained by the absolute difference value of two frames. It can reflect the movements in the frames. Supposing the background of video is static, if the difference is below some value, it means there is no movement and it may be caused by noise. Otherwise, it caused by movement. This value is called threshold. So moving pixels and static background pixels can be distinguished by thresholding \( T \) in (2). And we call this image mask. The value of \( T \) is determined in part B of II.
\[
Z_{k,k-1}(i,j) = \begin{cases} 
1, & \text{if } FD_{k,k-1}(i,j) > T \\
0, & \text{else}
\end{cases}
\]  

(2)

**B. Morphological Operations**

Morphological operations are usually performed on binary images where the pixel values are either 0 or 1. And they play a key role in applications such as machine vision and automatic object detection. There are three primary morphological functions: erosion, dilation, and hit-or-miss. Others are special cases of these primary operations or are cascaded applications of them. In this paper, morphological operations are used to obtain continuous moving regions. In order to remove holes and isolated pixels of the mask, morphological operations, such as DILATE, FILL, CLOSE, are performed in sequence. And morphological FILL operation is especially necessary.

**C. Moving Regions**

Moving pixels or background ones obtained by one frame difference image are always mistaken for others because of occluded or uncovered situations. So Twice Frame Difference was proposed in [5], which can be defined by the logic AND operation of two frame difference images. In order to make all moving pixels continuous and filter isolated pixels, moving regions are obtained by morphological CLOSE operation.

**D. Moving Region Combination**

Moving regions obtained in part C of III always include some isolated or small regions. And a moving object is always divided into several isolated regions. So combining moving regions is necessary, that is to say, two or more regions will be combined to one if they are near enough. Some people use distance of center points of regions to judge whether combining or not. In this article, we combine two or more regions to one when the nearest distance of them is below a threshold. Before calculating the nearest distance \( D \) of two isolated regions (such as region A and B), we first divide surroundings of region A into eight parts, which can be reduced to three classes (I, II, III, seen in Fig.1-1). In this article, relationship of region A and B will include three cases. Case 1 means that region B is in area I of region A completely (seen in Fig.1-2). Case 2 means that region B is above (or below) region A, and may be completely or partly in area II of region A (seen in Fig.1-3). Case 3 means that region B is in the left (or right) of region A, and may be completely or partly in area III of region A (seen in Fig.1-4). So we will define the nearest distance of two regions according to different cases.

Positions of region A and B are including above three situations. And the nearest distance \( D \) is calculated as following:

\[
D = \begin{cases} 
\sqrt{D_x^2 + D_y^2}, & \text{if case1} \\
\min(|y_B(l) - y_A(2)|, |y_A(l) - y_B(2)|), & \text{if case 2} \\
\min(|x_B(l) - x_A(2)|, |x_A(l) - x_B(2)|), & \text{if case 3}
\end{cases}
\]  

(3)

When \( D \) is below a threshold, region B will be combined into region A. And the top left corner and bottom right corner of region A will be updated accordingly.

![Figure 1. Classification of region A surroundings](image)

**E. Label Moving Regions**

After combining near regions, we obtain bigger regions. Now, what we do is to label left regions with rectangles.

**IV. EXPERIMENT RESULTS**

**A. Morphological FILL Operation**

Morphological operations can be used to remove isolated pixels and holes in the images. Especially morphological FILL operation is necessary, which is helpful for removing holes of mask images. It can affect directly the continuousness of moving regions.
B. Moving Region Combination

Definition of the nearest distance stated in part D of III is useful and efficient to combine neighbor moving regions. And it does better than combing by distance of center points. From Fig.3, it can be seen that moving regions can not be combined though their distance is very near when combined by distance of center points, since the distance of their center points is above the threshold (Fig.3-1), while the results combined by nearest distance is different (Fig.3-2).

C. Moving Objects Detection Results

Based on the proposed moving objects detection algorithm, the experiments are made and the results are shown in Fig.4. We label different moving objects with different color rectangles. When moving object appears and its region area is above a threshold, it can be considered as moving object and will be labeled out with rectangles.

Figure 2. Use of morphological FILL operations

Figure 4. Moving objects detection results of different video sequences

V. CONCLUSION

In this paper, we proposed an automatic moving object detection algorithm based on frame difference and region combination. First of all, we proposed an automatic threshold selection method to obtain moving regions of frames. After that, region combination was made successfully according to the proposed definition of the nearest distance. The experimental results show the proposed algorithm is efficient and automatic in moving objects detection for video surveillance applications.

ACKNOWLEDGMENT

This work is supported by the Natural Science Foundation of Shandong Province (Y2007G62, Y2007G44, Y2007G43) and the Outstanding Young and Middle-aged Scientists Foundation of Shandong Province (2008BS01019).

REFERENCES