

Motion Detection Based on Twice Segmentation

Jiangkai Dong¹, Yule Yuan², Yong Zhao³

Abstract. Motion detection is widely used in computer intelligent visual processing such as object tracking, object recognition and behavioural analysis. In this paper, Motion vector is adopted in order to get the accurate rough motion region. Motion vector can remove the small false motion objects. Frame difference method is conducted to get the motion matrix, and seed fill algorithm can be used to get the connectivity of the motion matrix. We can get the motion energy in the meantime. Some obvious false samples can be filtered out based on the motion vector. Considering of the regional characteristic of the motion objects in practical situation, we can set a size threshold to exclude these false motion regions. What's more, the motion energy can be used as the measurement of motion detection in the subsequent refine segmentation stage. Twice segmentation is needed to obtain the accurate motion object. This method can be efficient with less noisy point and better object skeleton with shadow removal. To confirm the effectiveness of the proposed method, we tested the method in different scene and background environment.

Keywords: Motion detection; Motion vector; GMM; Twice segmentation; Shadow removal

¹Jiangkai Dong

Key Lab. of Integrated Microsystem Sci. & Eng. Applications, Peking University Shenzhen Graduate School, China
e-mail: dongjiangkai@126.com

²Yule Yuan

Key Lab. of Integrated Microsystem Sci. & Eng. Applications, Peking University Shenzhen Graduate School, China
e-mail: yuanyl@pkusz.edu.cn

³Yong Zhao (✉)

Key Lab. of Integrated Microsystem Sci. & Eng. Applications, Peking University Shenzhen Graduate School, China
e-mail: yongzhao@pkusz.edu.cn

1 Introduction

Motion detection is a critical procedure in most video process. As the first phase of video analysis, it is important in motion tracking, motion recognition or human-computer interaction. Generally, motion detection can be divided into two parts: background segmentation and motion detection. Many methods have been proposed to solve this problem. Image frame difference is the mostly used methods. The method makes use of the pixel difference between the sequential frames in the video frame flow. Usually, we choose two or three frame difference. This algorithm has a high speed which can be used in real-time scene and dynamic background. However, it is sensitive to illumination changing and more likely to detect holes which split the motion objects. The other one is the background difference which utilizes the pixel difference between the current frame and an updating background frame. Its efficiency depends on a good background model and the updating strategy. But it performs badly when the background is dynamic changing. The other one is the optical flow method. It refers to the instantaneous speed of the motion pixels in which the iterative analysis is needed. This operation is time-consuming and inaccurate, so it has a bad real-time performance.

To overcome these drawbacks, various methods have been proposed. Plenty of algorithms have been used to model the background such as Histogram, LBP or GMM. The Mixture of Gaussians Models (GMM) is used to model some complex backgrounds [1]. In GMM, each pixel has its correspondent multiple mixture models. GMM is sensitive to small noisy point and the illumination changing. Local Binary Pattern (LBP) is an effective texture description operator which is used to measure and extract the local image texture information. LBP is presented as a binary string, which describes the texture information in a pixel region [2 3]. LBP is invariant to illumination and perform much better than GMM, but it failed to reflect the changing over time on this local descriptor. In this paper, we proposed an efficient method based on motion vector using the twice segmentation. We get the motion vector which removes the noisy block at the block level rather than the pixel level. Then we refine the rough motion objects by using an improved background difference model according to the motion vector and its energy. At last, shadow removal based on chromaticity and texture feature and morphological method can help to polish the refined motion objects.

The rest of this paper is organized as below. Section 2 introduces the motion vector and the detailed procedure of motion detection such as image blocking, motion vector, motion refinement and shadow removal. Section 3 demonstrates the experimental results in different test videos. In section 4, we conclude the entire paper.

2 Motion Detection

In this section, we need to introduce the specific techniques. Firstly, motion vector is obtained upon the blocked image using the method of frame difference. GMM background model is used to refine the rough motion vector. Most background modelling methods cannot remove the shadows which may cause the false detection and object adhesion. In this paper, a method combining the colour and texture information is used to remove shadows [4].

2.1 Motion Vector

Motion vector, imitating the concept of math, is a structure containing the connectivity, position and the pixel variance of this motion region. The connectivity implies the number of the adjacent blocks which also help to decrease the holes and the faked mini-motion. The position locates the exact point in the whole frame, which is presented as a rectangle. The pixel variance can be a measurement to update the background frame. The motion vector is defined as below:

$$mv[1, \dots, K] = (x_0, y_0, x_1, y_1, E, S) \quad (2.1)$$

Where, K is the number of the motion vectors in the frame, (x_0, y_0, x_1, y_1) is the position of the motion vector, E is the energy and S is the size of the vector.

Considering of the dynamic changing background, frame difference is a good method to get the motion vector. In order to decrease the computation, the video frame must be partitioned. We use the block level as the compute unit instead of single pixel [5]. The pixel of the block is presented by the local extreme maximum pixel value or the mean value. After blocking the frame, we define $M[i, j]$, $1 \leq i \leq M$, $1 \leq j \leq N$ as the motion matrix. Then the M can be defined as follow:

$$M[i, j] = \begin{cases} |s[m, n]|, & \text{if } \max_{i \leq m \leq i+8, j \leq n \leq j+8} |s[m, n]| > th1 \\ 0, & \text{otherwise} \end{cases} \quad (2.2)$$

Where, $s[m, n]$ is the pixel difference of the pixel block, and $th1$ is the threshold for motion detection. In addition, the maximum of the pixel difference is saved in M . Motion vector can be obtained upon the motion matrix. The most important operation is to get the connectivity of the motion blocks. As showed in the Fig.2.1, There are 30 motion blocks in the frame, and we need to compute the connectivity of each motion block. In this example, the connectivity is 1,3,6,7. The connectivity of each block is the size of the motion vector. Here we can set a threshold in order to remove the faked motion samples such as vectors dyeing with the dark blue and the baby blue. There are only three vector in this frame can be the positive motion object such as vectors dyeing with green, red and yellow. The adjacent blocks with the same right connectivity can be combined as a new motion vector.

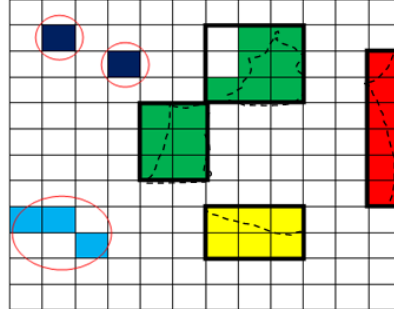


Fig. 2.1 Motion vectors in the blocking frame. The left three motion regions are removed because of its connectivity.

The motion vector can be represented as a rectangle. And energy E is the quadratic sum of the motion blocks in this motion vector which can be defined as below:

$$mv[k].E = \sum_{t=1}^{mv[k].S} M[i, j]_t^2 \quad (2.3)$$

2.2 Twice segmentation

We use the twice segmentation based on the motion vector to detect the motion. Firstly, we get the motion vector which can be used to update the background image when we refine the rough segmentation frame. In this paper we choose the GMM as the background modeling. The overflow of the whole method is showed as Fig.2.2:

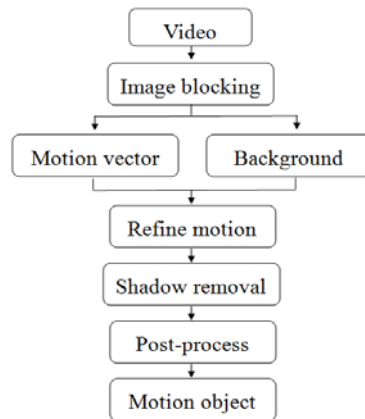


Fig. 2.2 Overflow chart of the proposed twice segmentation method. As showed in the fig.2.2, the most important part is the motion vector.

2.2.1 Rough Motion Region Segmentation

Rough motion region segmentation has many advantages. It decreases the computation, and the motion vector can remove the tiny noisy holes. Frame difference based on the blocking frame is used to obtain the motion vector. Then, seed fill algorithm is performed on the frame difference image in order to get the connectivity of motion vector. Energy is calculated for each motion vector.

2.2.2 Refine Motion Region Segmentation

As showed in Fig.2.1, the motion vector is rectangular. The edge of rough motion regions is inaccurate. The paper proposes a method based on the improved GMM background difference to refine the edge pixel of the motion vector. Unlike the traditional background difference, subtraction can only performs on the local region just as the dashed line region showed in Fig.2.1, which can largely reduce the complexity.

Let $b_t[i, j]$ be the background model obtained by GMM, $f_t[i, j]$ be the current frame in time t , and $d_f[i, j]$ be the background difference. Similarly, a threshold is set to decide whether the pixel belongs to the motion object or the background. $d_f[i, j]$ is defined as below:

$$d_f[i, j] = \begin{cases} \text{abs}(f_t[i, j] - b_t[i, j]) & \text{if pixel}[i, j] \in mv[k] \text{ and } mv[k].E > th2 \\ 0, & \text{otherwise} \end{cases} \quad (2.4)$$

$$\text{motion}[i, j] = \begin{cases} \text{yes,} & \text{if } d_f[i, j] > th3 \\ \text{no,} & \text{otherwise} \end{cases} \quad (2.5)$$

The threshold is the empirical value according to the video scene and the illumination condition.

2.2.3 Shadow Removal

The refined motion object contains the faked object pixels. Shadow is easy classified as the motion object because shadow shares the same intensity change and texture features. Judging from the features of removing the shadow, all the methods can be divided into four categories: chromaticity, geometry, physical and texture [4]. In this paper we choose the chromaticity and magnitude texture features to remove the shadow. HSV color feature can be effective in making a distinction between foreground and shadow [6]. Assumed block b is the shadow region, the color shadow model can be formed as below:

$$vLthres \leq \frac{F_b^v}{B_b^v} \leq vHthres \quad (2.6)$$

$$|F_b^h - B_b^h| < hthres \quad (2.7)$$

$$F_b^s - B_b^s < sthres \quad (2.8)$$

Where, F_b is the chromaticity value in the current frame while B_b is the same value in the background image. h, s , and v are the corresponding color components. But the chromaticity is sensitive to the illumination, so texture feature is used to help removing the candidate shadows. For each pixel block, gradient is computed as the texture, which shares the same texture with the foreground object [7]. Then the texture difference can be used to divide the shadow and the foreground.

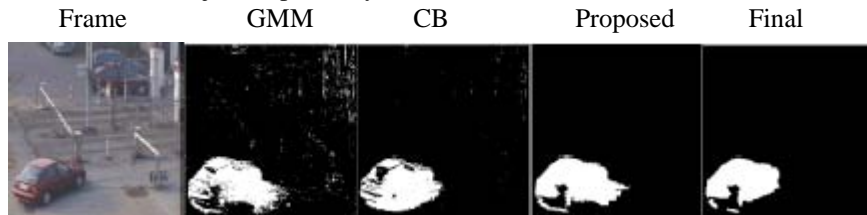
$$\nabla_b = \sqrt{\nabla_x^2 + \nabla_y^2} \quad (2.9)$$

$$\alpha_b = \arctan(\nabla_y / \nabla_x) \quad (2.10)$$

The block b , we get the maximum value of the gradient of each block. First of all, the gradient texture both in the shadow and the background should be remarkable compared to the threshold. The gradient ratio and the angle difference are calculated. The two values are used to count up the number of pixels sharing the similar gradient texture, which is the measurement of the gradient correlation.

3 Experimental Results

This section presents the experimental results. In this paper, we compared 3 methods: GMM, CB and the proposed method. Different videos are used to test the algorithms. As showed in Fig.3.1, the left column is the raw video frames, and the others are the results of GMM, CB [8], our proposed without shadow removal and the final motion object respectively from the second column to the last one.



(a) The 58th frame in the scene of the campus



(b) The 190th frame in the scene of Hallway



(c) The 159th frame in the scene of highway1



(d) The 378th frame in scene of IndoorGTest1



(e) The 280th frame in scene of intelligent room



(f) The 229th frame in scene of Laboratory



(g) The 60th frame in scene of tree

Fig. 3.1 Experimental results of GMM, CB, our method and the final results with the shadow removal in different video.

Compared to GMM and CB, the proposed method combined the GMM and the motion vector which considers the motion size and motion energy. This can help remove the small noisy points produced by the GMM. The last column is the result of shadow removal. In video c, this method can remove the most part of the shadow. Compared to GMM, the proposed method can remove most of the noisy

points by using the motion vector at the cost of less computation than other methods such as connected component method, and it has less ghost holes than CB. Our proposed method segments a smooth and integrated motion object region which makes it easier for the subsequent shadow removal.

4 Conclusions

This paper presented an effective and efficient strategy for motion detection. The method is sensitive to small disturbance such as illumination and gentle breeze in dynamic background. The improvement of the two methods, average difference and background difference, is much simpler and efficient. Experimental results indicate that the proposed method tends to produce the motion object with fewer holes and more accurate edge, being able to handle the outdoor illumination change. This method is simple and can be used in different scene such as vehicle detection, indoor intruder and so on. As we all know, most applications are all used in specific hardware platform. So the motion detection is the simpler the better. This method costs less memory, which can be used in RAM or other embedded system.

5 References

1. Stauffer C., Grimson W. E. L. (2000) Learning patterns of activity using real-time tracking. *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, 22(8), 747-757
2. Heikkilä M., Pietikainen M. (2006) A texture-based method for modeling the background and detecting moving objects. *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, 28(4), 657-662
3. Shimada A., Taniguchi R. I. (2009). Hybrid background model using spatial-temporal LBP. In *Advanced Video and Signal Based Surveillance*, 2009. AVSS'09. Sixth IEEE International Conference on (pp. 19-24). IEEE
4. Sanin A., Sanderson C., Lovell B. C. (2012). Shadow detection: A survey and comparative evaluation of recent methods. *Pattern recognition*, 45(4), 1684-1695
5. Zhao Y., Gong H., Lin L., Jia Y. (2008). Spatio-temporal patches for night background modeling by subspace learning. In *Pattern Recognition*, 2008. ICPR 2008. 19th International Conference on (pp. 1-4). IEEE
6. Cucchiara R., Grana C., Piccardi M., Prati A. (2003). Detecting moving objects, ghosts, and shadows in video streams. *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, 25(10), 1337-1342
7. Sanin A., Sanderson C., Lovell B. C. (2010). Improved shadow removal for robust person tracking in surveillance scenarios. In *Pattern Recognition (ICPR)*, 2010 20th International Conference on (pp. 141-144). IEEE
8. Kim K., Chalidabhongse T. H., Harwood D., Davis L. (2005). Real-time foreground-background segmentation using codebook model. *Real-time imaging*, 11(3), 172-185