Advances in Engineering Research, volume 114
International Conference on Advances in Materials, Machinery, Electrical Engineering (AMMEE 2017)

Moving object detection in video sequences

Di Ren [a] and Bing Xu [b]

Shanghai Institute of Technology, Shanghai 200000, China

[a]huoyushizhe@163.com, [b]xubing@163.com

Keywords: Moving object detection, Adaptive Gauss mixture background model, local texture features, YCbCr color features, Shadow detection.

Abstract. Moving object detection belongs to the primary processing stage of visual analysis, as well as the premise and basis of object tracking and behavior analysis. Foreground motion images containing shadows are obtained by an adaptive Gauss mixture background model. To better remove shadows and get moving objects, this paper has proposed a shadow detection algorithm based on local texture feature and YCbCr color features to detect shadow images, then, shadow images are subtracted from foreground motion images, and the final moving object is obtained by filtering and contour filling, which better achieves the detection of moving objects in video sequences.

1. Introduction

The intelligent video surveillance is a broad field for the current research and application. The intelligent video analysis includes the detection, identification and tracking. The detection of the moving object is the premise and foundation for the achievement of video analysis, its fastness and accurateness has a great impact on the subsequent analysis. The existing methods of moving object detection are roughly divided into: optical flow method[1], frame differential method[2] and background subtraction[3]. Of which the background subtraction becomes the most commonly used detection algorithm for the current time by the high detection accuracy and the low computational complexity[4].

Background subtraction is to calculate the mean of pixels by pixel counting on several frame data. Subtraction is conducted on current frame and background image in video sequences, Pixel points with small difference are regarded as background area, while Pixel points with large difference are regarded as movement region. Gauss mixture background model based on color feature adopts multiple mixture distributions to maintain each pixel point, it can better overcome the change of background environment disturbance, meanwhile, it will also mistake shadows as moving objects. The shadow detection based on texture feature has better robustness to light changes, but it is also easy to occur detection holes to moving objects with less texture information, causing the detected motion object to be incomplete[5].

By the background subtraction, this article utilizes OpenCV to establish the adaptive Gaussian mixture background model and get the moving images of prospect; conduct the shadow detection and get the images of shadow with the association of local texture features and YCbCr color features; the moving images of prospect subtract the images of shadow for the final moving objects through the filter and the outline.

2. Gauss Mixture Background Model

In video sequences, each pixel is represented by a plurality of Gauss models, for a point $P(x_0, y_0)$ in sequence image, if the observed value at t time is x_t, then, a series of observations are $\{x_1, x_2, x_3 ... x_t\}$, it is regarded as an independent statistical stochastic process, and described by K Gauss mixture distributions, the probability distribution of point $P(x_0, y_0)$ at the time t is gained from following formula[6]:

$$P(x_t) = \sum_{i=1}^{K} w_{i,t} \times N(x_t, \mu_{i,t}, \Sigma_{i,t}) \quad (1)$$

$w_{i,t}$ is the weight of i mixed Gauss distribution at time t. $\mu_{i,t}$ and $\Sigma_{i,t}$ are the mean and covariance matrices of i distribution, N is Gaussian probability-density function:

$$N(x_t, \mu_t, \Sigma_t) = \frac{1}{(2\pi)^{\frac{\pi}{2}} \left|\Sigma_t\right|^{\frac{1}{2}}} e^{-\frac{1}{2}(x_t-\mu_t)^T \Sigma_t^{-1}(x_t-\mu_t)} \quad (2)$$

$$\Sigma_{i,t} = \sigma_i^2 \cdot I \quad (3)$$

K is the number of Gaussian distribution in the hybrid model, generally the value range is 3 to 5, σ_i and μ_i are the variance and mean of the Gauss distribution model in the dimension of i [7], I is the unitary matrix.

Arrange the Gaussian Models in the number of K from large to small in accordance with w/σ ,take the previous models in the number of B as the background model:

$$B = \arg\min_b \left(\sum_{k=1}^{b} w_k > T\right) \quad (4)$$

In the equation, T is the threshold value of weight function.

Get a frame of new image, if each of the current pixel value of this image and the difference value of the background model meet the equation of $\left|X_t - \mu_{i,t-1}\right| \le \lambda\sigma_{i,t-1}$. Then it should be considered that this pixel X_t matches the Gaussian model. In the equation, λ is the matching constant, generally it takes 2.5. Update the respective parameters of this model:

$$\begin{cases} w_{i,t} = (1-\alpha)w_{i,t-1} + \alpha \\ \mu_{i,t} = (1-\rho)\mu_{i,t-1} + \rho X_t \\ \sigma_{i,t}^2 = (1-\rho)\sigma_{i,t-1}^2 + \rho(X_t - \mu_{i,t})^2 \end{cases} \quad (5)$$

In the equation, $\alpha(0 \le \alpha \le 1)$ is the learning rate, it determines the update rate of the background model parameters, if the value of α is larger, the upgrading is faster, but in order to reduce noise, usually, α is taken a smaller value. ρ is parameter update rate, $\rho \approx \alpha / w_{i,t}$. If the current pixel value X_t does not matched with any one of the Gauss models, then the new distribution will replace the minimum weight Gauss distribution, and the standard deviation will take the larger value, and the weight is the same as the weight of the replaced Gauss distribution. Because new distribution is added, the ownership value should be normalized, and each distribution is rearranged from large to small according to w/σ, then according to the above steps, X_t will be determined whether it is moving foreground pixels or background pixels. Gauss mixed background model can effectively overcome the interference caused by the illumination change and the disturbance caused by periodic perturbation of background image.

3. Shadow Detection

Under the lighting condition, the shadow will move with the moving object for which it must be considered in the image process for the elimination of it, otherwise the accurateness for the separation of images will be reduced. This article conducts the shadow detection with the use of the local texture features and the YCbCr color features respectively.

3.1 The Original LBP Operator

The original LBP operator (OLBP) is that within a window area of 3x3, with the central pixel as the threshold value, conduct the finite difference with the pixels that are neighbors within the area respectively, if the neighbored pixel is larger than the central pixel, then conduct the binarization, mark the neighbored pixel as 1, otherwise it should be 0. Mark all the neighbored pixels within the area, and the local binary pattern can be gained. Reflect the texture feature value of this area through the code value of LBP in the mean shift of the following equation.

$$LBP(x_c, y_c) = \sum_{n=0}^{n-1} 2^n \mathrm{H}(i_n - i_c) \tag{6}$$

In the equation, (x_c, y_c) is region center pixel, i_n is the brightness value of n pixel in the region, i_c is the brightness value of the central pixel.

The function H is defined as follows:

$$\mathrm{H}(x)=\begin{cases}1 & x \geq 0 \\ 0 & x < 0\end{cases} \tag{7}$$

3.2 Improve LBP Operator

Because the performance of the original LBP operator to resist the sampling noise is not strong[4], improve the function $\mathrm{H}(x)$ as $\mathrm{H}(x)=\mathrm{H}(|i_n - i_c| - T)$, of which T is the threshold value, it can be known through the multiple times of experiment, if the threshold value is small, it will create much noises, if the threshold value is too large, the LBP texture features will be insufficient, this article takes 6 for T.

3.3 Shadow Detection Based on Texture

Shadow detection is carried out according to the texture similarity between the shadow area and the corresponding position in the background model[8]. Texture similarity is defined as follows[9]:

$$S_M(x, y) = 1 - \frac{M_{LBP}^{C}(x, y) - M_{LBP}^{B}(x, y)}{255} \tag{8}$$

In the equation, $M_{LBP}^{C}(x, y)$ and $M_{LBP}^{B}(x, y)$ respectively present texture eigenvalues of the current frame and point (x, y) in the background frame, the more similar the texture values of these two, the value of $S_M(x, y)$ closer to 1; conversely, it is closer to 0. In this paper, a similarity judgment threshold $T1 = 0.8$ is set, if the value of $S_M(x, y)$ is larger than and equal to 0.8, then, the pixel point (x, y) is determined as shadow; conversely, it belongs to moving object.

3.4 Ycbcr Color Space

YCbCr is a kind of color space, based on the human eye to the brightness of this feature. Of which Y represents the luminance component, Cb and Cr are respectively represent the blue and red chrominance component.

The comparison expression transferred from the RGM color space to the YCbCr color space is as follows:

$$\begin{cases} Y = 0.299 \times R + 0.587 \times G + 0.114 \times B \\ Cb = (B - Y) \times 0.564 + 128 \\ Cr = (R - Y) \times 0.713 + 128 \end{cases} \tag{9}$$

3.5 Shadow Detection based on Ycbcr Color Space

The brightness of the pixel in the foreground shadow is lower than the brightness of the pixel in the background region, but the hue is similar to the hue of the pixel in the background area, this serves as a basis for shadow judgment. Shadow detection is performed on YCbCr color space, for the correlation of each component is high and the calculation is large in RGB color space, while, luminance component Y and chrominance component Cb, Cr in YCbCr color space are independent with each other[10], which can reduce the probability of false detection of shadow detection under illumination and other changes. According to the shadow judgment basis, when the pixel point (x, y) satisfies the following formula, it is regarded as shadow:

$$\begin{cases} \left|C^{Y}(x,y)-B^{Y}(x,y)\right| \le T_{Y} \\ \left|C^{Cb}(x,y)-B^{Cb}(x,y)\right| \le T_{Cb} \\ \left|C^{Cr}(x,y)-B^{Cr}(x,y)\right| \le T_{Cr} \end{cases} \tag{10}$$

In the equation, $C^{Y}(x,y)$, $C^{Cb}(x,y)$, $C^{Cr}(x,y)$ are the three component values of the image point of prospect (x,y), $B^{Y}(x,y)$, $B^{Cb}(x,y)$, $B^{Cr}(x,y)$ are the three component values for the corresponding points of the background image, T_Y, T_{Cb}, T_{Cr} are the threshold values for the judgment of shadow of the three component values which are Y, Cb, Cr respectively. Through multiple times of experiments, this article sets T_Y, T_{Cb}, T_{Cr} as 30, 12, 12 respectively. Reference [10], during the seek for the solution of $C^{Y}(x,y)$, $C^{Cb}(x,y)$, $C^{Cr}(x,y)$, $B^{Y}(x,y)$, $B^{Cb}(x,y)$, $B^{Cr}(x,y)$, utilize the thinking of field, conduct the average calculation for each of the pixel points and the other pixel points in its 8 fields, take the average value as the pixel value of this point which is similar to the mean filter, but this article has conducted the Gaussian filter during the pre-process for the image, and it calculates the pixel value of each pixel points by the weighted average calculation which is better for the calculation effect than that of the single average calculation, therefore, it is not conducted with the mean processing here.

4. Moving Object Detection Process

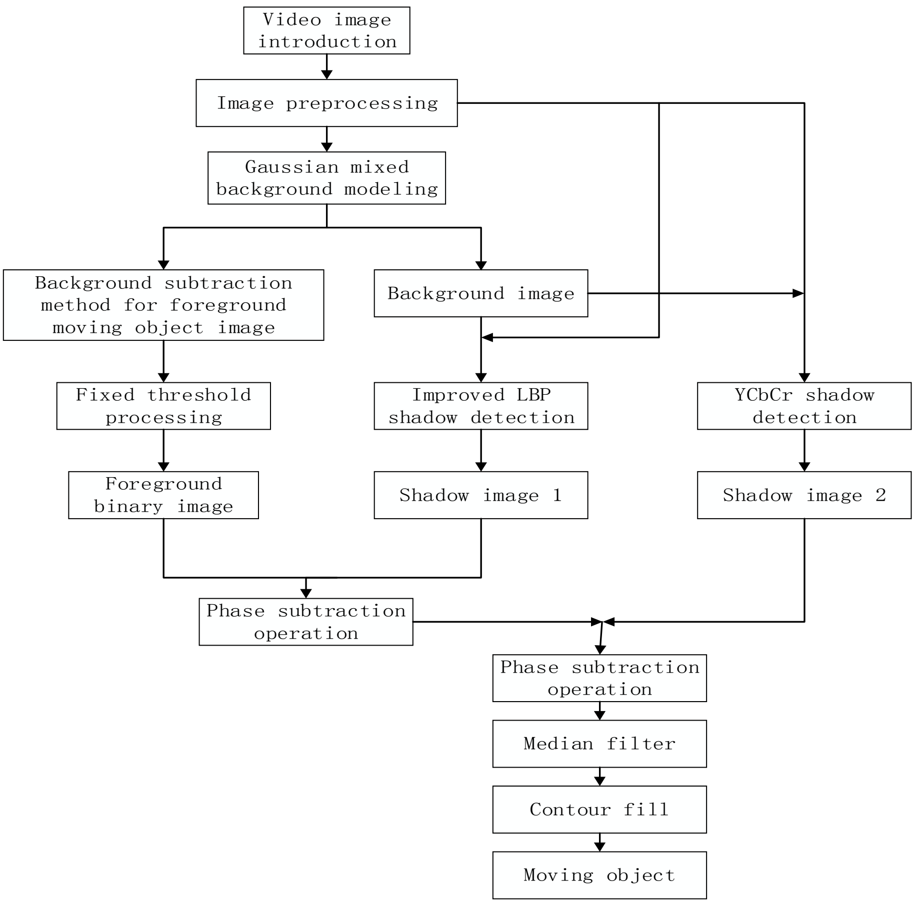

Figure 1. Moving object detection process

When realizing moving object detection, this paper firstly adopts bilinear interpolation method to zoom the image to the right size (depending on the size of the incoming video), this is to speed up the later image processing, and then Gauss filter processing is conducted, reference [11] used median filtering for image preprocessing, but after many experiments, it was found that the median filter was less effective than the Gauss filter, therefore, Gauss filtering is used to process the image. An adaptive Gauss mixture model is used to obtain foreground motion images (including large amounts of Shadows) and background images. In this paper, we propose an improved LBP shadow detection method to obtain the shadow image of the original image and the background image 1, then foreground motion images are used to subtract them to obtain a moving object image with less shadow 1. Meanwhile, the original image and background image are detected by YCbCr shadow to get shadow image 2, subtract the shadow image 1 from the previous moving object image 2, then the median filter is used to obtain the final moving object.

5. Experimental Results and Analysis

The following experimental environments are all Visual Studio 2012 and OpenCV3.1, the equipments are all configured with Intel Core i7-6700HQ Quad-core processor, the basic frequency is 2.6 GHz, 64-bit operating system. There are two groups for the test videos, one group is the indoor scene, the other group is the outdoor scene, both are the mainstream test videos. Get the moving object that contains a lot of shadows through the adaptive Gaussian Mixture Background Model, with the integration of the final moving objects gained by the shadow removal method suggested in this article, the experimental results is indicated as the below image.

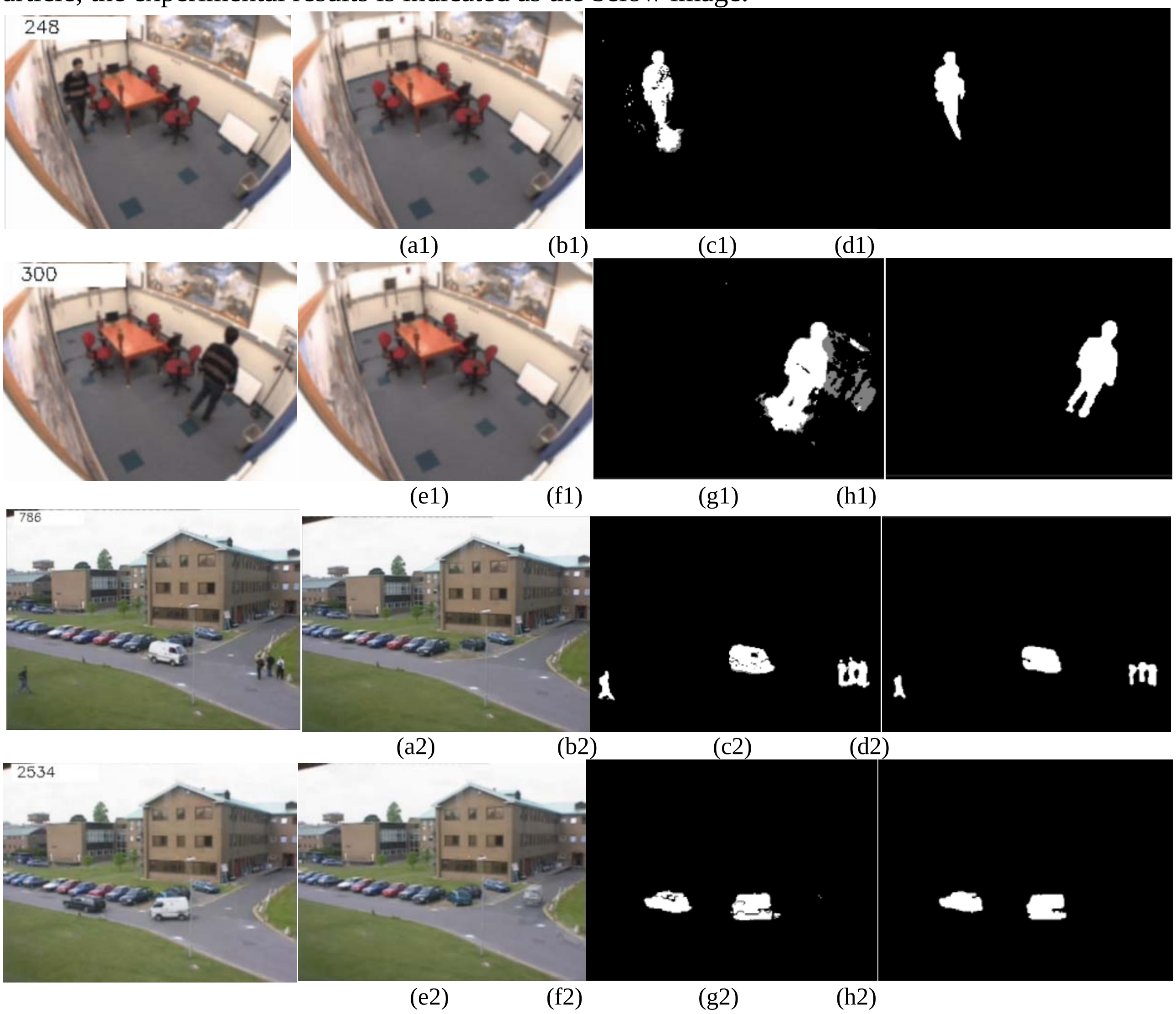

Figure 2. Experimental results in video sequences

(a1)The raw image in the No. 248 frame;(b1)The background image in the No. 248 frame;(c1)The GMM algorithm image in the No. 248 frame;(d1)Our proposed method image in the No. 248 frame;(e1)The raw image in the No. 300 frame;(f1)The background image in the No. 300

frame;(g1)The GMM algorithm image in the No. 300 frame;(h1)Our proposed method image in the No. 300 frame;(a2)The raw image in the No. 786 frame;(b2)The background image in the No. 786 frame;(c2)The GMM algorithm image in the No. 786 frame;(d2)Our proposed method image in the No. 786 frame;(e2)The raw image in the No. 2534 frame;(f2)The background image in the No. 2534 frame;(g2)The GMM algorithm image in the No. 2534 frame;(h2)Our proposed method image in the No. 2534 frame

Indicated in the above images of(d1),(h1),(d2),(h2), the algorithm in this article can better detect the moving objects in the sequence of videos compared with the algorithm of GMM, the outline is clear, the background is clean and it can effectively eliminate the shadows, although there is the existence of emptiness inside the moving objects, most of the objected outline exist with it, it does not impact the work of identification and tracking at later stage. In addition, as it is indicated in (b1),(f1),(b2),(f2), the adaptive background images are all gained through the Gaussian mixture background modelling which quitely match the actual background and may prove that the adaptive Gaussian mixture model plays the function well in the detection of the moving objects.

6. Summary

This article establishes the adaptive Gaussian Mixture Model and gains the moving images of prospect and the background model through OpenCV. In order to eliminate the shadow and get the better moving objects, the joint of texture features and color features is suggested to conduct the detection of the shadows, and then utilize the moving images of prospect gained before to subtract the images of shadows, get the final moving objects throguh the medium filtering and the outline filling. Although the algorithm in this article can better detect the moving objects, under the conditon when the object color is very close to the background color, the effect for the detection of this algorithm needs still to be improved and needs to solve this problem in the later research.

References

[1]. Guowu YUAN, Zhiqiang CHEN, Jian GONG, et al. A moving object detection algorithm combining optical flow method and three frame difference method [J]. Journal of Chinese Mini-Micro Computer Systems. Vol. 34 (2013) No. 3, p. 668-671.

[2]. Baoyuan CHEN, Zhichao HUO, Guangyi CHEN, et al. An improved three frame differential moving object detection algorithm[J]. Applied Science and Technology. Vol. 43 (2016) No. 2, p. 10-13.

[3]. Yu SHEN, Xinxin WANG. Video moving object detection method based on background subtraction and interframe difference method[J]. Automation and Instrumentation. (2017) No. 4, p. 122-124.

[4]. Shunfei WANG, Junhua Yan, Zhigang WANG. Improved method of moving object detection based on local joint feature[J]. Chinese Journal of Scientific Instrument. Vol. 36 (2015) No. 10, p. 2241-2248.

[5]. Yonghua JIA, Wencong Zhang, Shiliang PU. Background modeling method based on texture feature and color feature fusion[J]. Chinese public security. (2011) No. 4, p. 156-160.

[6]. Jianquan YANG. Research on object detection and shadow removal algorithm in video surveillance (Master, National University of Defense Technology, China, 2006). p. 23.

[7]. [Wei ZHU, Chunguang ZHAO, Leqing OU, et al. OpenCV image processing programming examples [M]. Publishing House of Electronics Industry, 2016, p. 346.

[8]. Ling ZHANG, Yimin CHENG, Shiming GE, et al. Moving shadow detection method based on texture [J]. Opto-Electronics Engineering. Vol. 35 (2008) No. 1, p. 80-84.

[9]. Ying DING, Wenhui LI, Jingtao FAN, et al. Moving object detection algorithm based on Choquet fuzzy integral [J]. Chinese Journal of Electronics. Vol. 38 (2010) No. 2, p. 263-268.

[10]. Ying ZHENG, Juan ZHAO, Su XIAO. Shadow removal method based on YCbCr color space and local texture weighting[J]. Computer knowledge and technology. Vol. 9 (2013) No. 32, p. 7349-7354.

[11]. Tongxue ZHOU, Ming ZHU. Moving object detection in video images[J]. Chinese Journal of Liquid Crystals and Displays. Vol. 32 (2017) No. 1, p. 40-47.