# Moving Object tracking Based on Background Subtraction Combined Temporal Difference

Ssu-Wei Chen, Luke K. Wang, Jen-Hong Lan

***Abstract*—**In this paper, we propose a new tracking method that uses Three Temporal Difference(TTD) and the Gaussian Mixture Model(GMM) approach for object tracking. TTD method is the use of a continuous image subtraction. The GMM approach consists of three different Gaussian distributions, the average, standard deviation and weight respectively. There are two important steps to establish the background for model, and background updates which separate the foreground and background . This paper combines the TTD and GMM object tracking. The advantages of TTD quick calculations and the disadvantage is a lack of complete object tracking. The advantage of GMM is complete results of the operation the disadvantage is not a complete object tracking, GMM result of the operation complete but disadvantages include computing for a long time with more noise. These two methods can complement each other and image filtering results in the successful tracking of objects.

***Keywords*—**Background Model, Temporal Differencing, Gaussian Mixture Model, Object Tracking.

## I. Introduction

VIDEO based moving object tracking such as visual surveillance, medical image processing, face tracking, object tracking , image recognition, etc. is one of the important missions in the computer vision field. In recent years, the new object tracking algorithms have been proposed, object tracking algorithms are often used include shift, optical flow, Kalman filter[1] etc. Shift algorithm has good recognition rates and high accuracy, but if the camera lens or tracking object is moving too fast, this will lead to an inaccurate algorithm[2]. Due to the optical flows large algorithm size, it is generally better to use a simple one for background tracking[3]. Based on the Kalman filter algorithm itself it cannot be applied in a nonlinear system lead to decrease accuracy. In this paper, we combine GMM and TTD method, GMM can be used in the context of a complex environment, the TTD is used by the time difference of three consecutive images, and by adjusting the threshold to extract an image of the movement area, for more complex background is not easy to track moving objects, the inner image will produce an empty phenomenon, it's effect alone is not a complete algorithm. Currently, the research literature offers many kinds of background models, the traditional Gaussian mixture model uses a mixture of the K Gaussian model distribution of background pixels that will be the scene to establish the statistical model for each pixel, and periodically update the background[4]. The Gaussian model based on the screen to pure only when the background images with each pixel to calculate the mean and standard deviation , and the future prospects of the background information as a basis for classification[5][6][7][8][9]. Another method is to calculate the probability of classification through the foreground and background pixels using the Bayesian classifier, or with the probability of hidden Markov model to achieve the classification of foreground and background[10]. Fig.1 shows the basis process of the algorithm.

The rest of this paper is organized as follows. section 2 describes the TTD and morphological method with median filter. Section 3 describes the GMM model contains background models and update method. Section 4 describes our proposed method. Section 5 describes experimental results. Finally, section 6 presents our conclusions.

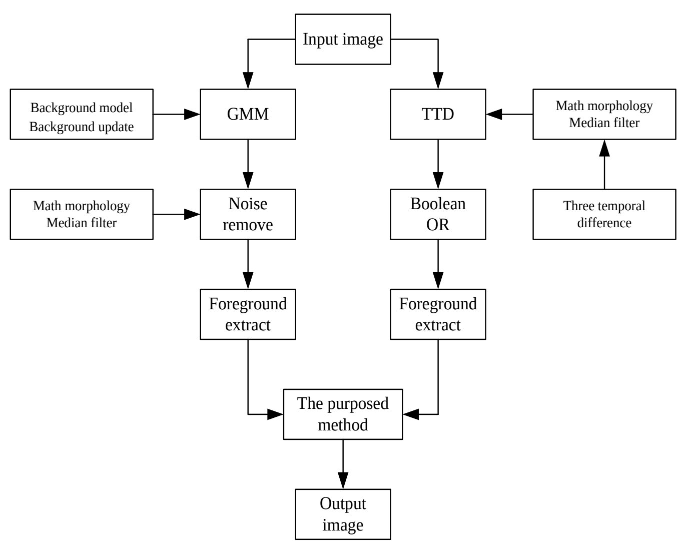



Fig. 1 Flow chart of the proposed object tracking

Ssu-Wei Chen is with the Department of Electrical Engineering of National Kaohsiung University of Applied Sciences, Kaohsiung, 80778 Taiwan, R.O.C (corresponding author to provide phone:+886-6-5931668; fax:+886-6-5932265 ; e-mail: 1098404109@cc.kuas.edu.tw ).
Luke K. Wang is with the Department of Electrical Engineering of National Kaohsiung University of Applied Sciences, Kaohsiung, 80778 Taiwan, R.O.C (e-mail: lwang@mail.ee.kuas.edu.tw ).
Jen-Hong Lan is with the Department of Electrical Engineering of National Kaohsiung University of Applied Sciences, Kaohsiung, 80778 Taiwan, R.O.C (e-mail: obiwan@adm.cgmh.org.tw ).

## II. Three Temporal Differencing Method

### A. Three Temporal differencing

TTD is the principle of continuously time subtracting image pixels. The traditional method will cause the internal cavity TD case, thus the moving object shape is not complete for the follow up tracking and identifying moving objects will not be able to provide complete information. The traditional image subtraction method is determined by subtracting the previous image from the current image to obtain motion information, this paper uses three consecutive image subtraction methods, and then uses logic algorithm segmentation motion blocks, if the three successive images were $I_{n-1}(i,j), I_n(i,j), I_{n+1}(i,j)$, the mathematics is as follows:

$$I_A(i,j) = |I_{n-1}(i,j) - I_n(i,j)| \quad (1)$$

$$I_B(i,j) = |I_n(i,j) - I_{n+1}(i,j)| \quad (2)$$

$$I_C(i,j) = \begin{cases} 1 & I_A(i,j) \cup I_B(i,j) \geq th' \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

To obtain the $I_C(i,j)$, we give a threshold, this threshold can remove noise , and can be set for different light conditions, then we set the threshold conditions for 20.

### B. Mathematical Morphology

Morphological erosion and dilation in the image processing is an important foundation. Dilation is the image of the object's mathematical computing size, the dilate is a collection of the operations defined.

$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \emptyset\} \quad (4)$$

Where $\emptyset$ the empty set and B as structural elements. The erosion is the image of objects smaller or thinner, erosion and dilation similar to the mathematical definition.

$$A \ominus B = \{z | (B)_z \cap A^c \neq \emptyset\} \quad (5)$$

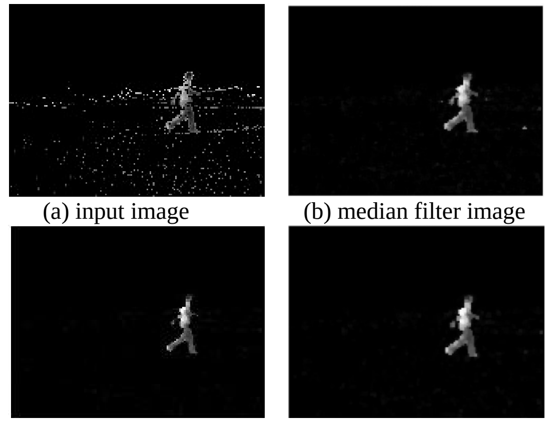
Erosion of A by B is a structural element of the origin of all the set positions, in which translation of the B and A's background does not overlap.

### C. Median filtering

Median filter is a commonly used image process, we need to reduce noise before image processing, median filter algorithms determine the principles of an odd pixel window W, window size of each pixel arranged according to Gray, middle gray value instead of the original F(i,j) the gray value, gray value as the center of the window g(i,j).

$$g(i,j) = \text{median}\{F(i-k, j-l), (k, l \in W)\} \quad (6)$$

Where W is the selected window size, F(i-k,j-l) for the window W of the pixel gray value, usually an odd number of pixels in the window. Figure 2 shows the median filter and morphological erosion and dilation results, this combination provides us with excellent results.



(a) input image (b) median filter image

(c) erosion image (d) dilation image

Fig. 2 Integrated noise filtering algorithms

## III. Gaussian Mixture Model

The Gaussian mixture model is a single extension of the Gaussian probability density function. As the GMM can approximate any smooth shape of the density distribution, so often used in image processing in recent years for good results. Assuming the Gaussian mixture model consists of and the combination of Gaussian probability density function, the Gaussian probability density function of each has its own mean, standard deviation, and weight, the weights can be interpreted by the corresponding Gaussian model of the frequency, they more often appear in the Gaussian model the higher the weight. The higher frequency of occurrence, then find the maximum weight on the Gaussian probability density function, Finally, the Gaussian probability density function of the means pixel value is background image.

### A. Background model

Background subtraction is one of the most common method of object segmentation, this process contains two steps: background and update model. The basic theory of the Gaussian mixture model is as long as the number of Gaussian mixtures, an arbitrary distribution can be in any of the precision is mixed with a weighted average of these Gaussian approximation.

$$P(X_t) = \sum_{i=1}^{K} \omega_{i,t} * \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (7)$$

where K is the number of distributions , $\omega_{i,t}$ is an estimate of the weight, $\mu_{i,t}$ is the mean value , $\Sigma_{i,t}$ is the covariance matrix , and where $\eta$ is a Gaussian probability density function

$$\eta(X_t, \mu_{i,t}, \Sigma_{i,t}) = \frac{1}{(2\pi)^{\frac{\pi}{2}} |\Sigma|^{\frac{\pi}{2}}} e^{-\frac{1}{2}(X_t - \mu_t)^T \Sigma^{-1} (X_t - \mu_t)} \quad (8)$$

these parameters determine the characteristics of this density function, such as the center of the function shape, width and direction and so on.

### *B. Background update*

The known algorithms, if not updated, the step operation time will be very long, we must use the iterative method to update the mean, standard deviation and the weight to reduce the time required. New steps before must set the basic parameters, as we select the number of Gaussian components C=3,Number of background components are M=3, Positive deviation threshold D=2.5, learning rate between 0 to 1, in this paper we set α =0.05,foreground threshold =0.4. The gain weight ,mean and standard deviation are updated by using(9)-(12)

$$\rho = \frac{\alpha}{W_{i,j,k}} \tag{9}$$

$$W_{i,j,k} = (1-\alpha)W_{i,j,k} + \alpha \tag{10}$$

$$\mu_{i,j,k} = (1-\rho)_{i,j,k} + \rho * I_n \tag{11}$$

$$\sigma_{i,j,k} = \mathrm{sqrt}\left[(1-\rho)\sigma^2_{i,j,k} + \rho(I_n - \mu_{i,j,k})^2\right] \tag{12}$$

Although the computational complexity of the GMM is high, but it can provide better results. If new entrants cannot be matched to any pixel of a Gaussian probability density function, update the pixel value of mean, then initialize the weights and the standard deviation.

## IV. The Proposed Method

Combine the GMM with the TTD method we can obtain the results of two different images, we have to combine the advantages of GMM and TTD. There are two methods that will be combined using the logical AND function in the past literature, conventional TD algorithms led to an internal cavity, and then use a logical AND with the GMM method. There is still emptiness, according to the light in different conditions are prone to this problem, therefore, we must use a method to fill the image enhancement GMM. Figure 3 shows the $F_{GMM}(x,y)$ coordinates of the location markers to identify the object The coordinates of the image as the main control location, in the $F_{TDD}(x,y)$ method to identify the location of the object, voids will occur in some situations, $F_{GMM}(x,y)$ coordinates of the location and $F_{TDD}(x,y)$ coordinates of the location overlap, $F_{TDD}(x,y)$ coordinates of the location corresponding to the $F_{GMM}(x,y)$ coordinates on the object, then the object will be tracked by fragments corresponding to $F_{GMM}(x,y)$ coordinates of the objects position.

$$F_{GMTD}(x,y) = \begin{cases} F_{TDD}(x,y) & F_{GMM}(x,y) = 0 \\ F_{GMM}(x,y) & F_{GMM}(x,y) \geq 0 \end{cases} \tag{13}$$

We use the GMM approach as the main algorithm and the other algorithms as auxiliary, generated by the main algorithm to meet the emptiness. There is not only one way to fill the hole produced by the main algorithm, different algorithm can be used to fill the void phenomenon supporting the main algorithm, using the stacking method, the stacking of different algorithms to achieve the perfect effect.

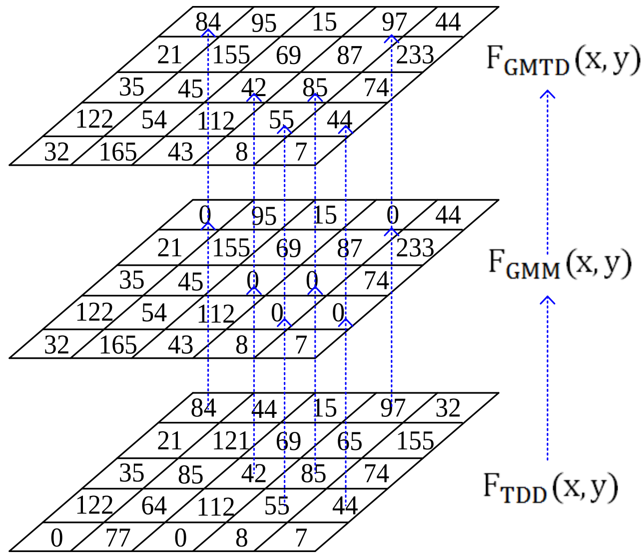



Fig. 3 The proposed method

## V. Experimental And Results

In this section, we show experimental results of the proposed object tracking method. The proposed algorithm was implemented in MATLAB 7.6(2008a), and tested in windows XP SP3 with Intel dual core I5 CPU with a memory of 6GB.

The object video sequences come from Stanford Vision Laboratory which is publicly available, the size of the video sequences is 240×320 pixels. Figure 4 show that pedestrian tracking results, figure 5 show that vehicle tracking results, figure 6 we show many different types of filter algorithms, the results showed that the proposed method is better than other methods.

(a)The 125th frame

(b)The 250th frame

(c) The 500th frame

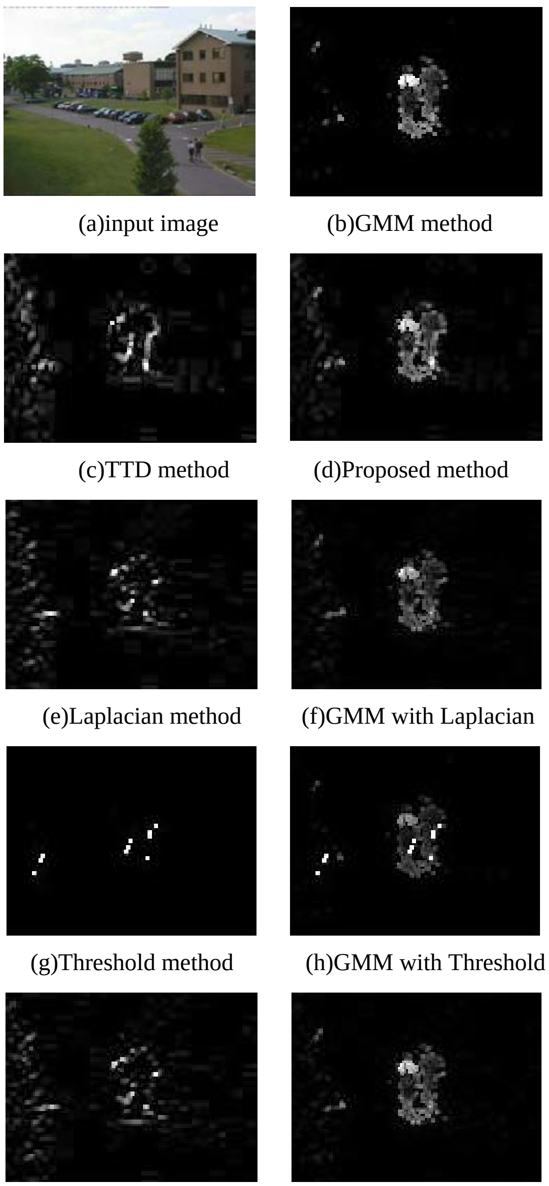
(d)The 700th frame

Fig. 4 Pedestrian tracking results

(a)The 1565th frame (b)The 1610th frame

(c)The 1635th frame (d)The 1800th frame

Fig. 5 Vehicle tracking results



(a)input image (b)GMM method

(c)TTD method (d)Proposed method

(e)Laplacian method (f)GMM with Laplacian

(g)Threshold method (h)GMM with Threshold

(i) Sobel method (j)GMM with Sobel

Fig. 6 The result of various filters

## VI. Conclusion

In this paper, we proposed the GMM and TTD method successfully applied in a continuous image. We used the GMM approach as the main tracking algorithm, with morphological and median filtering to remove noise. The success of the foreground and background segmentation and found the object coordinates. On the other hand, we used the TTD method to subtract successive images, also using morphological and median filters to remove noise. Due to TTD method tracking object incompletely, supporting the main algorithm GMM, the proposed method there is still room for improvement. We can replace the TTD method, fill the emptiness produced by the phenomenon of GMM, but the best algorithms in the future still need more testing to be able to get the perfect result.

## References


[1] Comaniciu D, Ramish V, Meer P. Real-time tracking of non-rigid objects using mean shift ,IEEE Conference Computer Vision and Pattern Recognition,2000:142-149.

[2] Barron J L,Fleet D J, Beauchemin S S. Performance of optical flow techniques. International Journal of Computer Vision, 1994, 12(1):43-77.

[3] Li P H, Zhang T W. Unscented Kalman filter for visual curve tracking. Image and Vision Computing,2004,22(2):157-164.

[4] C.Stauffer, W.E.L Grimson. "Adaptive Background Mixture Models for Real-Time Tracking," in Proc. Computer Vision and Pattern Recognition Conf.,vol. 2, Fort Collins, CO, USA, June 1999,pp.246-252.

[5] Sohail Nadimi and Bir Bhanu "Physical Models for Moving Shadow and object detection in video",IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.26,Aug 2004,pp 1079-1087.

[6] Tao Zhao and Ram Nevatia, "Tracking Multiple Humans in Complex Situations", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 26,Sep 2004,pp 1208-1221.

[7] Ahmed Elgammal, Ramani Duraiswami , David. Harwood, and Larry Davis, "Background and foreground modeling using nonparametric kernel density estmation for visual suveillance." IEEE Transactions on Pattern Abalysis and Machine Intelligence,vol. 90, July 2002 ,pp 1151-1163.

[8] Rita Cucchiara, Massimo Piccardi, Andrea Prati, "Detection Moving object, Ghoust, and Shadows in Video Streams", IEEE Transactions on Pattern Abalysis and Machine Intelligence,vol. 25, Oct 2003 ,pp 1337-1342.

[9] Wenmiao Lu and Yap-Peng Tan,"A Color Histogram Based People Tracking System",IEEE International Symposium on Circuits and System, vol.2, May 2001, pp137-140

[10] Leonid Sigal , Stan Sclaroff and Vassilis Athitsos , "Skin Color-Based Video Segmentation under Time-Varying Illumination", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 25 July 2004,pp 862-877.