Research Journal of Mathematics and Statistics 5(4): 43-47, 2013
ISSN: 2042-2024, e-ISSN: 2040-7505

Submitted: September 24, 2013 Accepted: October 30, 2013 Published: November 30, 2013

Moving Object Tracking based on Background Extraction Using Mean Algorithm and Three Temporal Difference Algorithm

[1, 2]Ahmed Mahgoub Ahmed Talab, [1]Zhangcan Huang
[1]School of Sciences, Wuhan University of Technology, Wuhan, China
[2]College of Engineering, Elimam ELMahdi University, Kosti, Sudan

Abstract: In this study, we propose a new tracking method that uses Three Temporal Difference (TTD) and the Mean Algorithm (MA) to approach the tracking of an object. TTD method is used for continuous image subtraction while the MA method is used for the extraction of Background image. The proposed method was compared with different methods used in the field; the comparison clearly shows that the method is reliable, quickly and precise. This method has the advantage that it is fast and successfully tracks the objects and extract background image, also no shadow and noise was associated with the application of the method. Experimental work shows that our method is improved relatively to the other widely used techniques.

Keywords: Background extraction, mean algorithm, object tracking, surveillance, three difference algorithm

INTRODUCTION

Extracting the background from a video sequence is an open problem with very practical applications ranging from camera surveillance systems to human-computer interactions (Bazzani *et al*., 2009; Yang *et al*., 1998). Intelligent Transportation System can be defined as means of transport realize the modernization of Technology Traffic Management Information, such as information technology, data and communication technology, automatic control technology, computer Processing (Hernández *et al*., 2002); With the fast development of video technology, detecting moving objects from image sequences is a crucial step in Intelligent Transportation System. The efficiency and accuracy of continuous image processing depend on the results of moving objects detecting (Meijin *et al*., 2009). Background extraction is a crucial step in many automatic video applications (Barnich and Van Droogenbroeck, 2009). These new methods allow researchers to begin modeling real world processes under varying conditions (Stauffer and Grimson, 1999). Video based moving object tracking in the computer vision field such as visual surveillance, medical image processing, face tracking, object tracking; the image recognition is one of the significant tasks in computer vision applications. Recently, researchers developed new object tracking algorithms; object tracking algorithms are often used to include shift, optical flow, Kalman's filter (Comaniciu *et al*., 2000). Shift algorithm is featured with a proper recognition rate and considerable accuracy. Inaccurate algorithm is associated when the camera lens or tracking object is moving too fast (Barron *et al*., 1994). It is recommended to use a simple algorithm for background tracking because of the size of the optical flows large algorithm (Li *et al*., 2004). Based on the Kalman filter; it cannot be applied in a nonlinear system that will lead to decrease accuracy (Chen *et al*., 2011).

In this study, Three Temporal Difference (TTD) and Mean Algorithm (MA) are combined. Mean Algorithm (MA) had an excellent performance in stable environment and complicated environment; the advantage of this method is its successful extraction of background image. In the case of environment in short time had a big change the mean algorithm yields poor performance. Three Temporal Differences (TTD) have not a background and insensitive in a situation at night and sunshine TTD has a good performance than the mean algorithm. TTD is used by the time difference for three consecutive images by adjusting the threshold to extract an image of the moving area. For more complicated background, it is hard to track the moving objects. So the combination advantage of mean algorithm and three temporal differences is to produce good performance.

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 (Sigal *et al*., 2004).

ALGORITHMS

Three temporal differencing: TTD is the rule of continuously time subtracting image pixels. The formal method will cause the internal cavity TTD case, thus

Corresponding Author: Ahmed Mahgoub Ahmed Talab, School of Sciences, Wuhan University of Technology, Wuhan, China

the moving object shape is not achieved for the follow-up tracking and identifying moving objects will not be able to provide complete information (Chen *et al*., 2011). The conventional image subtraction method is specified by subtracting the previous image from the current image to take out motion information; this study uses three consecutive image subtraction methods and then uses logic algorithm segmentation motion blocks, if the three successive images were $I_{n-1}(x, y)$, $I_n(x, y)$, $I_{n+1}(x, y)$, the mathematics is as follows:

$$I_A(x, y) = |I_{n-1}(x, y) - I_n(x, y)| \tag{1}$$

$$I_B(x, y) = |I_n(x, y) - I_{n+1}(x, y)| \tag{2}$$

$$I_C(x,y) = \begin{cases} 1 & I_A(x,y) \cup I_B(x,y) \geq T \\ 0 & \text{Otherwise} \end{cases} \tag{3}$$

To obtain the $I_C(x, y)$, we give a threshold, this threshold can remove noise and can be set for different light conditions, then we set the threshold conditions for 10.

Mean algorithm: In mean algorithm the proposed possibility of the pixel being a background is greater than being foreground. In a very small space of time (e.g., frames), the grey level of background will change within a very small range, but the grey level of the foreground objects is vary with each car. The gray color variation within the car is existed and it depends on each part color. Grounded on this assumption, frequently certain value of pixel can be extracted and categorized to be a background image.

This algorithm utilizes in 2-D image sequences, for every pixel (x, y), the corresponding point's values in former N frames are: $I_{t-N}(x, y), \ldots, I_{t-2}(x, y), I_{t-1}(x, y)$ compute the sequence of values by applies the mean algorithm and adopt the result as background value of current image, the computing formula of background magnitude is:

$$B(x,y) = Mean(I_{t-N}(x,y), I_{t-N+1}, \ldots, I_{t-2}(x,y), I_{t-1}(x,y))$$

This study proposed a new method to obtain the value of unchanging background by employing three temporal difference algorithms and the changed background pixels are estimated through mean algorithm. The result from the new method is approximate to real background.

THE PROPOSED METHOD

The objective of this study is to propose a new method to obtain background and foreground image from moving video, because most of the current methods are suffering from slowness of image extraction, the extracted image most of the time featured by shadow, in addition all other methods need more space memory devices. The proposed method used the Mean Algorithm (MA) with the Three Temporal Difference algorithm (TTD) then we can obtain the results of three different images; and we have to combine the advantages of MA and TTD. Hereinafter we outline the main steps:

Calculation of the background in this method: For the proposed method if the number of frames is less than 30 we use the average image to compute the initial background. If the number of frames is greater than or equal to 30 the learning rate (α) is used to update the background image computation; the following computing formula is used:

$$BG_{n+1}(x,y) = (1-\alpha) \times BG_n(x,y) + \alpha \times I_{n+1} \tag{4}$$

$BG_{n+1}(x, y)$ = New background pixel at position (x, y)
$BG_n(x, y)$ = Old background pixel at position (x, y)
$I_{n+1}(x, y)$ = The pixel at position (x, y) of the new image
α = Learning rate

This method can really eliminate the deficiency in mean method.

Calculation of the foreground in this method:

- Three sequential frames separated by fixed interval, $I_{n+1}(x, y)$, $I_n(x, y)$ and $I_{n-1}(x, y)$, respectively, were converted in to Gray images for simple and real time.
- Using $I_{n+1}(x, y)$, $I_n(x, y)$ and $I_{n-1}(x, y)$
Through (1) we can get two difference image $d_n(x, y)$, $d_{n+1}(x, y)$.

$$d_n(x,y) = \left| I_n(x,y) - I_{n-1}(x,y) \right|$$

$$d_{n+1}(x,y) = \left| I_{n+1}(x,y) - I_n(x,y) \right|$$

- Comparing the value pixels in $d_n(x, y)$, $d_{n+1}(x, y)$ to certain threshold (T_2) we determine the foreground according to Eq. (3):

$$fg = \begin{cases} 1 & \text{if } (d_n > T_2 \,\&\, d_{n+1} > T_2) \\ 0 & \text{Otherwise} \end{cases}$$

- The connect components in computed foreground were found and then the components with size less than certain threshold were removed.

Through the steps above, the background and foreground can be extracted accurately.

Comparison of standard of algorithms: The performance of the proposed algorithm is achieved by using the root Mean Square Error (RMS) to assess if the background is better or not and the average time is used to asses if the algorithm is fast enough or not. We should find a pure background frame STBG (x, y) from the video or man-made it for the comparison. The RMS error and average time were defined as follow:

$$RMS_{error} = \frac{1}{M \times N} \sqrt{\sum_{x=1}^{M} \sum_{y=1}^{N} (BG(x,y) - STBG(x,y))^2} \tag{5}$$

$$\text{average time (t)} = \frac{1}{t} \sum_{i=1}^{t} (time_i) \tag{6}$$

$time_i$ = The time of the ith frame take

The smaller the RMS error, the higher is the efficiency of the background image. The smaller RMS_{error} imply that the pixel of background extraction is distributed closer to the real background scene. Conversely, it is worse. After judging background image; we use Eq. (7) to judge if the foreground image is better or not:

$$Bd = \left| background - frame_i \right| \tag{7}$$

Make Eq. (7) to use binarization where the white pixel is foreground. The white pixel is calculated by Eq. (7) and name COUNT is given to white pixel.

By Comparing the count obtained by different algorithms, if the algorithm has the small count and similar by a certain extent to the original image, then it is the best algorithm for background extraction.

EXPERIMENTAL AND RESULTS

In this section, we show experimental result of the proposed object tracking method. The proposed algorithm was implemented in MATLAB. (R2012 b) and tested in windows 8 with Intel (R) core (TM) i7-3632QM CPU @ 2.20 GHz 2.20 GHz with a memory of 4GB. The object video sequences come from MATLAB (traffic.mj2) which is publicly available; the size of the video sequences is 120 × 160 pixels Fig. 1. The video is clear and showed many different types of algorithm. Figure 1 showed that vehicle tracking results.

Based on the standard of comparison (5), (6) in above section, we can calculate the (RMS) and average time (t) corresponding to Fig. 1. The values are shown Table 1.

The results of experiments under the same conditions are shown Fig. 1. Figure 1b resulted from signal Gaussians and its quality is worse than in Fig. 1c.

(a) Original image (b) Siginal Guassian method

(c) Proposed method

Fig. 1: Background extraction based on different methods from another data

(a) Original image (b) TTD method

(c) Siginal Guassian method (d) Proposed method

Fig. 2: Foreground extraction based on different method and no use connect components

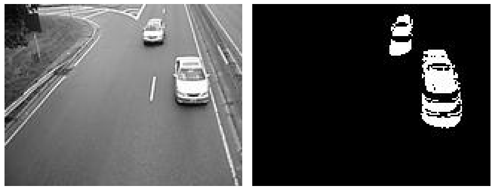
Table 1: The values of RMS and average time (t)

	Signal Gaussian method	Proposed method
RMS error	0.0478	0.0294
t	0.0509	0.0048

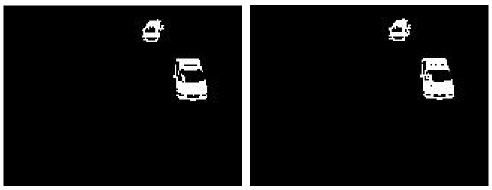
Table 2: The values of count through different algorithm

	TTD method	Signal Gaussians	Proposed method
Count	1851	19198	19199

The result of proposed background extraction showed in Fig. 1c and the moving car is almost eliminated in this image. The advantage of the proposed method is also proved in Table 1. The values of root-mean square error (RMS_{error}) of each algorithm in Table 1 showed that the coefficients of the new method are minimized. Thus, it is proven that the proposed method in this study has a better quality than others.



(a) Original image (b) TTD method



(c) Siginal Guassian method (d) Proposed method

Fig. 3: Foreground extraction base on different method and use connect components

Table 3: The values of count through different algorithm

	TTD method (b)	Signal Gaussian (c)	Proposed method (d)
Count	1823	860	1579

Based on the standard of comparison (7), we can calculate the Count (white pixels) and average time (t) corresponding to Fig. 2. The values is shown Table 2.

The results of experiments under the same conditions are shown Fig. 2. The Number of pixels for the extracted foreground image in video (traffic.mj2) from MATLAB was 19199 for the proposed method, while it was ranged from 1851 to 19199 for the other methods. This clearly showed that our method is better quality the number pixels in foreground image; because in the proposed method is no shadow.

Based on the standard of comparison (7) in above section, we can calculate the Count (white pixels) and average time (t) corresponding to Fig. 3. The values are shown Table 3.

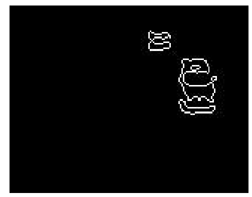
The results of experiments under the same conditions are shown in Fig. 3. The Number of pixels for the extracted foreground image in video (traffic.mj2) from MATLAB was 1579 for the proposed method, while it was ranged from 860 to 1823 for the other methods. This clearly showed that our method yields a better quality for foreground image; because no shadows were recorded while TTD method recorded some shadows.

Based on the standard of comparison (7) in above section, we can calculate the Count (white pixels) and average time (t) corresponding to Fig. 4. The values is shown Table 4.

The results of experiments under the same conditions are shown Fig. 4. The number of pixels for the extracted foreground image in video (traffic.mj2) from MATLAB was 3374 for mean with Laplacian method, while it was from 892 to 3374 for the other methods. This showed edge clearly that mean Laplacian method has a better quality than Laplacian method in foreground image; and in this method is no noise there.

Based on the standard of comparison (7) in above section, we can calculate the Count (white pixels) and average time (t) corresponding to Fig. 5. The values are shown Table 5.

(a) Original image (b) Laplacian method



(c) Mean with Laplacian

Fig. 4: Foreground extraction (edge) base on different method and use connect components

(a) Original image (b) Sobel method

(c) Mean with Sobel

Fig. 5: Foreground extraction (edge) base on different method and use connect components

Table 4: The values of count through different algorithm

	Laplacian method	Mean with Laplacian method
Count	892	3374

Table 5: The values of count through different algorithm

	Sobel method	Mean with Sobel method
Count	930	1043

The results of experiments under the same conditions are shown Fig. 5. The number of pixels of the extracted foreground image in video (traffic.mj2) from MATLAB was 1043 for the mean with Sobel method, while it was from 930 to 1043 for the other methods. This shows edge clearly that mean Sobel method has a better quality than Sobel method in the foreground image.

CONCLUSION

In this study, we propose Mean Algorithm (MA) and Three Temporal Difference algorithm (TTD) successfully applied in a continuous image. We used the mean method approach as the main tracking algorithm and the DTT method to subtract successive images, also used to connect components to remove noise. The experimental results of the proposed method in this study successfully extract background and foreground image. The experimental results yield that the proposed method is fast, accurate and no noise and shadow was associated with all extracted images.

ACKNOWLEDGMENT

The first author would like to thank all students in laboratory Prof Zhangcan Huang in Wuhan University of technology and gratefully appreciate Mr. Xiong Jiang for his assistance within the research.

REFERENCES

Barnich, O. and M. Van Droogenbroeck, 2009. ViBe: A powerful random technique to estimate the background in video sequences. Proceeding of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009), pp: 945-948.

Barron, J.L., D.J. Fleet and S.S. Beauchemin, 1994. Performance of optical flow techniques. Int. J. Comput. Vision, 12(1): 43-77.

Bazzani, L., D. Bloisi and V. Murino, 2009. A comparison of multi hypothesis kalman filter and particle filter for multi-target tracking. Proceeding of Performance Evaluation of Tracking and Surveillance Workshop at CVPR, pp: 47-54.

Chen, S.W., L.K. Wang and J.H. Lan, 2011. Moving Object tracking based on background subtraction combined temporal difference. Proceeding of International Conference on Emerging Trends in Computer and Image Processing (ICETCIP), Babgkok.

Comaniciu, D., V. Ramesh and P. Meer, 2000. Real-time tracking of non-rigid objects using mean shift. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp: 142-149.

Hernández, J.Z., S. Ossowski and A. Garcıa-Serrano, 2002. Multiagent architectures for intelligent traffic management systems. Transport. Res. C-Emer., 10(5): 473-506.

Li, P., T. Zhang and B. Ma, 2004. Unscented Kalman filter for visual curve tracking. Image Vision Comput., 22(2): 157-164.

Meijin, L., Z. Ying and H. Jiandeng, 2009. Video background extraction based on improved mode algorithm. Proceeding of 3rd International Conference on Genetic and Evolutionary Computing (WGEC'09), pp: 331-334.

Sigal, L., S. Sclaroff and V. Athitsos, 2004. Skin color-based video segmentation under time-varying illumination. IEEE T. Pattern Anal., 26(7): 862-877.

Stauffer, C. and W.E.L. Grimson, 1999. Adaptive background mixture models for real-time tracking. Proceeding of IEEE Computer Society Conference on Computer Vision and Pattern Recognition.

Yang, J., R. Stiefelhagen, U. Meier and A. Waibel, 1998. Visual tracking for multimodal human computer interaction. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '98), pp: 140-147.