Journal of Theoretical and Applied Information Technology
30th November 2013. Vol. 57 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

VIDEO-SURVEILLANCE SYSTEM FOR TRACKING A PERSON BY ESTIMATING THE TRAJECTORY

[1]BOUTAINA HDIOUD, [2]RACHID OULAD HAJ THAMI, [3]MOHAMMED EL HAJ TIRARI
[1,2] Mohamed V Souissi University (UM5S), National Higher School of Computer Science and Systems Analysis (ENSIAS), RIITM Research Group, Rabat, MOROCCO
[3]Insea, Rabat-Instituts, BP 6217, Rabat, MOROCCO
E-mail: [1]hdioud.boutaina@hotmail.fr, [2]rachid.ouladhajthami@gmail.com, [3]mtirari@hotmail.fr

ABSTRACT

This paper we propose a simple and efficient video surveillance system which detects and tracks a person in a video stream. Object detection is the most important and crucial step for any video surveillance system. In this paper, we separate foreground and background by using the statistical model of Gaussian Mixture (GMM). Interests points are identified in the detected regions (foreground) using the Harris detector. The analysis of connected components detected by subtracting the background allows grouped the pixels of moving objects in order to extract the center of gravity. Then, a boundary box is used to limit the area of connected components in order to detect the coordinates of the gravity center of the moving person. Finally, by projection in Euclidean plan we can get the trajectory of the person in motion and compute the Euclidean distance crossed in the scene.

Keywords: *GMM model, Connected Component, Boundary Box, Video analysis, Harris detector, Trajectory estimation.*

1. INTRODUCTION

For several decades, the video surveillance has grown tremendously, both in public places and sensitive areas within companies. Historically, the use of video surveillance was first involved the observation data by an agent, and recording videos to consult when needed. Video surveillance is to place surveillance cameras in a public or private place seeking to automatically identify events of interest in a variety of situations. Example applications include intrusion detection, activity monitoring, prevent theft, assault, fraud and manage incidents and crowd movements, and to ensure the security and traffic control. This control can help access various information and distinctive features of these people such as their trajectory.

Tracking people by their trajectories was objects of several researches. Many researchers have been proposed in this topic: J. Berclaz *et al.* [1] have been proposed a simple and robust method to track people with the overall trajectory model. They treated the paths individually and separately along the video sequence with the objective to avoid confusing people. F. Burkert *et al.* [2] suggested a model for tracking people based on sequences derived from aerial images of the camera systems installed on planes. The approach displays results of automatic detection and tracking of people from image sequences. Moreover, the trajectories of people are automatically interpreted through behavior in order to detect exceptional events. The trajectory of a person is important and useful information for many fields: security, marketing and other domains. For example, in commerce we use trajectories to analyze tendency of a customer's pathway. In [4], S. Takeda, *et al.* uses trajectories to track people using existing surveillance cameras based on the "Sensing web" framework. They would detect and track the people without putting their trajectories in its observation area.

Several experiments have shown techniques for detecting and tracking complex events. In the presence of a video stream to analyze, we are faced with a multitude of tasks to tackle: how to use preprocessing? how to segment and track moving areas? how to extract relevant information to draw the trajectory of detected objects? etc…. Each of these techniques has its own lock ups. For example, a crucial step for a proper analysis of the video is the correct detection of moving object.

The main task of a video sequences analyzing system is motion detection. Most motion detection algorithms presented in the literature [8, 9, 11 and 12] are presenting as methods of background subtraction. The basic operation of these methods is the separation of moving objects (Foreground) and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the static objects from the background. In this paper, we propose a new simple but beneficial idea for tracking people with their trajectories estimation in video using a single camera view in real time. The method proposed is based on three different steps: moving object detection, location of interest points on the moving object, tracking the object from estimation of trajectory.

In the following, we first present an overview of the proposed method. Each step of the method is detailed in sections 3 and 4. In section 5, we describe practical results and give frames to clarify each step. Finally, in section 6 is devoted to conclusion and perspectives.

2. OVERVIEW OF THE PROPOSED METHOD

Tracking people by estimating their trajectory has received much attention in the past few years, in all video surveillance systems; the method presented in this paper is based on three different steps: moving object detection, location of interest points on the moving object, tracking a person from an estimated trajectory. We firstly use background subtraction to detect objects of interest, this step is very useful as it permits the localization of moving objects in the image. There are many challenges in developing an efficient background subtraction algorithm. The most common prototype for performing background subtraction is to find an explicit model of background. Foreground objects (objects in motion) are then detected by calculating the difference between the current frame and this background model. In our case, the technique of GMM is commonly used. This method is most common in the case of background dynamic. It is very robust to sudden and gradual illumination changes, changes of position of objects in the background, etc... Then, foreground pixels detected by this technique are gathered in connected components. In the ideal case, one connected component corresponds to one object of the scene. Then, the Harris technique has been used so as to reduce the search space for tracking each moving object present in the scene. Once objects of interest are detected and tracked, we determine the displacement at any time in the scene by extraction of the Center of mass (centroid). Finally, we draw and calculate the trajectory traveled by each person.

The following diagram shows the steps used in our proposed people tracking system:

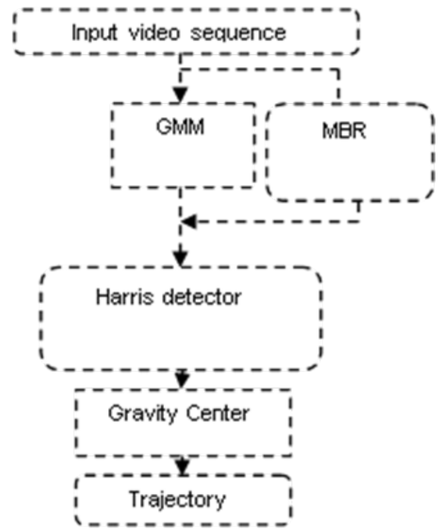

Figure 1. Proposed Steps Algorithm

3. MOVING OBJECT DETECTION

Moving object detection is the basic step for further analysis of video; it is a very important step because the following steps will be based on its results. Different techniques have been used in this stage, such as temporal background subtraction. The basic operation of these methods is the separation of moving objects (Foreground) and the static objects from the background, we present in this section the background subtraction method used in our system.

3.1 Foreground detection process

Foreground detection is an important pre-processing step for detecting the moving objects from the video. Pixels in the current frame that deviate significantly from the background are considered to be moving objects. These foreground pixels are further processed for object localization and tracking.

In our case we adopt the Gaussian Mixture Model for background subtraction [12]. It returns a foreground with minimal errors in terms of misclassified pixels. In this model each pixel location is represented by a number of Gaussian functions that sum up together to form a probability distribution. In the GMM formulation, each pixel $\{X_1, \ldots, X_t\}$, is modeled by K-Gaussians distributions. The probability of observing the current pixel value (pixel intensity variation) is:

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

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

with K is the number of distributions and η is a Gaussian probability density function:

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t-\mu_t)^T \Sigma^{-1}(X_t-\mu_t)} \quad (2)$$

μ_k, σ_k, and ω_k are estimated by the Expectation Maximization algorithm.

Figure 2. On The Left, The Original Frame, In The Middle, The Background And In The Right Extracted Foreground

4.1 Bounding box for interest moving people

In this step, we automatically surround the moving person detected in precedent step by using region props of the Matlab function and this by measuring the propriety bounding box of image regions, the minimal bounding box rectangle according to the following equation:

$$BBox = \{[x_{min}, y_{min}, x_{max}, y_{max}] \mid x, y \in O^i\}$$

where O^i denotes the set of the coordinates of points in the interest moving object; x_{max}, y_{max} is the left-top corner coordinates of the interest moving object and x_{max}, y_{max} is the right-bottom corner coordinates of the interest moving object.

Figure 3. In The Left: The Original Frame Of Foreground, In The Right: The MBR Surrounding Person

4. TRACKING PERSON FROM TRAJECTORY ESTIMATION

4.1 The Harris detector

After detecting the people in motion, Harris detector is applied for positioning the points of interest. Generally, points of interest are widely used in the literature for matching images: the notion of point of interest introduced by Moravec [5], to characterize where the signal is rich in information. According to Moravec [5], a point of interest is an image where the light intensity varies greatly in different directions. A lot of work has been done on the detection of points of interest [5, 6 and 7], and the detector is the most widely used Harris. This detector is based on the autocorrelation function.

We will measure the quality of a point using the Harris corner criteria, which is essentially a measure of the intensity variation in a neighborhood around the pixel.

If the relative coordinates of a pixel are given by X = [x, y]^T, then the quality of the point is measured by:

$$(X) = \det(G) + k(\mathrm{trace}(G))^2 \quad (3)$$

Where G is given by:

$$G = \begin{pmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_y I_x & \sum I_y^2 \end{pmatrix} \quad (4)$$

with respectively, I_x and I_y represent the horizontal and vertical intensity gradient at the pixel of interest, and then K is the scalar value determined by user.

Each tracked object is characterized by a set of points of interest. These points are tracked, frame by frame.

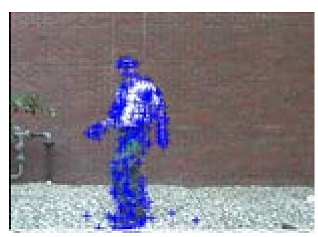

Figure 4. The Interest Points Detected By Harris

4.1 The center gravity of the interest points of moving people

After extracting the blobs (connected components), we determinate the centroid (Center of gravity) of each objects. Center of gravity [13] is defined as the unique point of the distribution of mass of interest points in space. For N interest points X_1, X_2…X_N with respectively I_1, I_2, …, I_N

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

their mass defined by the luminous intensities. The Barycenter is given by the equation

$$Br = \frac{\sum_{i=1}^{N} I_i \ X_i}{I_1 + I_2 + \ldots + I_N} \quad (5)$$

For the vector **Br** = (Br_1, Br_2, Br_3… Br_n) which is composed of the points Br_i with i is the number of gravity centers which is equal to the number of frames in the stream used.

Practically, the equation representing the segment line between two Barycenter displacements Br_1 (x_1, y_1) and Br_2 (x_2, y_2) is given by the following formula:

$$Y = \frac{y_2 - y_1}{x_2 - x_1} X + \left(y_1 - \frac{y_2 - y_1}{x_2 - x_1} x_1 \right) \quad (6)$$

with N points of interest give N-1 segments that are connected to trace the trajectory of the moving person.

4.1 Euclidean Distance

Euclidean distance between two points X and Y is defined as the length of the line segment connecting them.

In our digital image processing between two pixels P(x, y) and Q (u, v), the Euclidean distance function [13] is:

$$d_e(P, Q) = \sqrt{(x-u)^2 + (y-v)^2} \quad (7)$$

The length of traversed trajectory is given by the sum of N Euclidean distances.

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

To evaluate the performance of our method, we tested it on three real sequences taken from different base:

Video1: PETS2006 (Ninth International Workshop on performance evaluation of tracking and monitoring).

Video2: a standard video from web application

Video3: ATON sequence for laboratory and intelligent room.

The videos can be types of dynamic (changing illumination) or static background. The result shows the experimentation carried out in order to estimate the trajectory displacement of a person in motion.

The results are given and organized respectively in this order: The first test captured frame from our video sequence, second we surrounded foreground of person in motion with MBR, third we represented all those detected by Harris points of interest then we extracted Gravity center of connected components, finally Traversed trajectory forming by connecting gravity center of all components.

Figure 5. The Original Frame With Minimal Boundary Rectangle

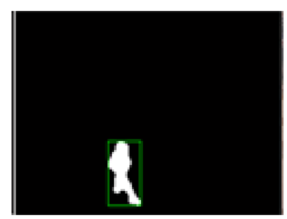

Figure 6. Extracted Foreground Surrounded by The MBR

Figure 7. The Gravity Center Of Connected Components

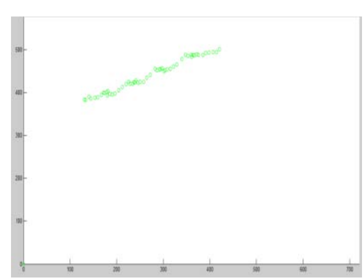

Figure 8. The Trajectory Of The Person In Motion (Video1)

Journal of Theoretical and Applied Information Technology
30th November 2013. Vol. 57 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The displacement vectors is given by $\mathbf{V}_1 = (v_1, \ldots v_n)$ where n is the number of frames in the video sequence which gives the traversed Euclidean distance d_1=786.5879

For the second sequence we have the following results:

Figure 9. The Original Frame With Minimal Boundary Rectangle

Figure 10. Extracted Foreground Surrounded By The Mbr

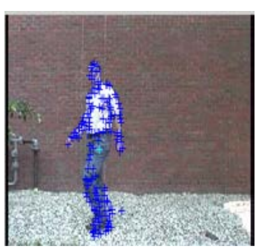

Figure 11. The Gravity Center Of Connected Components In MBR

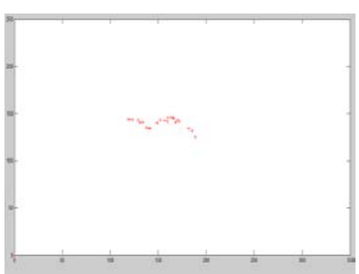

Figure12. The trajectory of the person in motion (video2)

The corresponding vector of displacement is given by $\mathbf{V}_2 = (v_1, \ldots, v_m)$ with m is the number of frames in the video 2. This test gives the following Euclidean distance: d_2= 524.3101

Finally, for our third sequence in test we have the following visual results:

Figure 13. The Original Frame With Minimal Boundary Rectangle

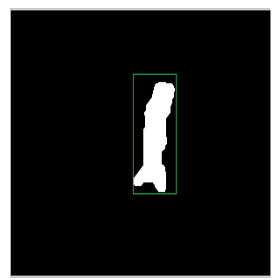

Figure 14. Extracted Foreground Surrounded By The MBR

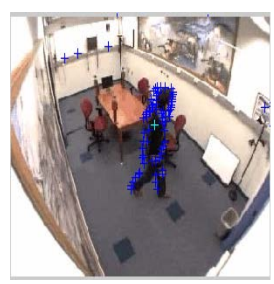

Figure 15. The Gravity Center Of Connected Components In MBR

As in the previous results of video1 and video2 the corresponding trajectory of the third video is:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

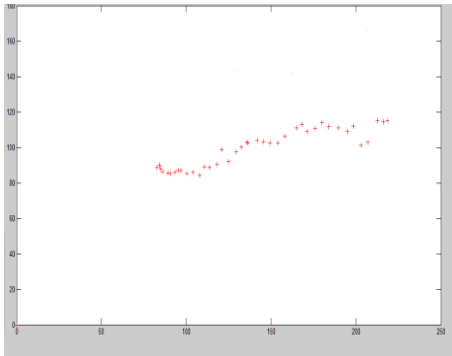

Figure 16. *The Trajectory Of The Person In Motion (Video*3)

And the sum of Euclidean distances gives the value: d_3=593.6067

As seen in the results, the calculated Euclidean distances corresponding to videos gives estimation to the displacement of the person in motion.

6. CONCLUSION AND FUTURE WORKS

In this paper, we proposed a new approach to achieve a video surveillance system for tracking a moving person detected by a static camera in a special area. Our approach is simple to implement compared to existing approaches. Practically, we can use this system for example to supervise a person in Hospitals, Banks and other environments, in order to analyze its behavior. In our system, the detection is performed in a simple way based on the GMM algorithm and the MBR (Minimum Boundary Rectangle) used to limit the search area of interest points in positioning the center of gravity of the connected components. Finally, by calculating the Euclidean distance of each person we have tracked the connected components and the points of interest detected by Harris detector. This helps to previously understand the behavior of the person in the area and supervise along the view by the camera, and gives an idea about the length of the crossed distance during the view and direction of the trajectory for tracking the person.

Our future work aims to analyze the profiles of the moving person in a store and generally in sale points and other interesting places. We will also seek to adapt the approach proposed in this paper for use in the detection and tracking of multiple people in video stream.

REFRENCES:

[1] J. Berclaz, F. Fleuret and P. Fua, “Robust people tracking with global trajectory tptimization” *EPFL - CVLAB* CH - 1015 Lausanne, Switzerland, 2006.

[2] F. Burkert, F. Schmidt, M. Butenuth and S. Hinz, ”People tracking and trajectory interpretation in aerial image sequences” Paparoditis N., Pierrot-Deseilligny M., Mallet C., Tournaire O. (Eds), *IAPRS, Vol. XXXVIII, Part* 3*A*, September 1-3, 2010, Saint-Mandé, France.

[3] A. Garcia-Martin and José M. Martinez, “On collaborative people detection and tracking in complex scenarios”, *Image and Vision Computing,* may 2012, 30 (4-5), pp. 345–354.

[4] S. Takeda, I. Mitsugami,, K. Kakusho, M. Minoh, "Obtaining Global People Trajectories Considering Tracking Failure in Each Camera", *International Workshop on Sensing Web* (in conjunction with the 19th International Conference on Pattern Recognition - ICPR2008), 2008.

[5] H. Moravec, “Towards automatic visual obstacle avoidance”. *In Proceedings of the 5th Int. Joint Conference on Articial Intelligence*, 1977, pp. 584. Cambridge, Massachusetts, USA.

[6] L. Kitchen and A. Rosenfeld. “Gray-level corner detection”, *Pattern Recognition Letters*, 1982, pp. 95-102.

[7] R. Deriche and G. Giraudon. “Acurate corner detetction: An analytic study”, *Technical Report* 1420, *INRIA Sophia-Antipolis*, France, 1991.

[8] P. Lalande, P. Bouthemy, “A statistical approach to the detection and tracking of moving objects in an image sequence”, *5th European Signal Processing Conference* EUSIPCO 90, 1990.

[9] L. Carminati, J. Pineau, M. Gelgon, “Human detection and tracking fro video surveillance applications in low density environment”, *SPIE VCIP’2003 SPIE 0277 -786X 5150*, 2003, pp. 51–60.

[10] D.H. Ballard and C.M. Brown. “Computer Vision”. *Prentice Hall Professional Technical Reference*, 1982.

[11] C. Wren, A. Azarbayejani, T. Darrell and A. Pentland. “Pfinder : Realtime tracking of the human body”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, 1997, pp. 780-785, 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

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

[13] Frank Y. Shih and Yi-Ta Wu, "Fast Euclidean distance transformation in two scans using a 3_3 neighborhood", *Computer Vision and Image Understanding 93*, 2004, pp. 195–205.