Journal of Theoretical and Applied Information Technology
10th February 2013. Vol. 48 No.1

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

AN EFFICIENT PEOPLE TRACKING SYSTEM USING THE FFT-CORRELATION AND OPTIMIZED SAD

[1]ABDERRAHMANE EZZAHOUT, [2]MOULAY YOUSSEF HADI AND [1]RACHID OULAD HAJ THAMI

[1]RIITM research group, ENSIAS, Souissi, Mohamed V University, Rabat, MOROCCO

[2]Faculty of Sciences, Ibn Tofail University, Kenitra, MOROCCO

E-Mail: [1]abderrahmane.ezzahout@um5s.net.ma, [2]hadi@gmail.com,[1]oulad@ensias.ma

ABSTRACT

The research questions which this study addresses are resumed in obtaining the picks of Optimized SAD and FFT-Correlation function. These values allow us to detect and track the blocs corresponding to the moving person. This Bloc-Matching system of our tracked bloc 8x8, simplify the process of detecting and tracking all the representative matrix of foreground. Practically, in this correspondence, we propose an implementation technique for detection and tracking system which detect and track people's in motion using a fixed single camera. These bi-techniques simplify the process of people tracking and give a new way for people tracking. Moreover, in this system we choose to detect moving people with the more robust algorithm (Gaussian Mixture Model) in term of misclassified pixels. Thereafter, we track a bloc 8x8 from the bounding box containing the moving person by returning the position of the similar bloc in the next frame. This bloc position is detected by maximization of the Fast Fourier Transform-Correlation (FFT-Corr) and minimizing the error: optimized Sum of Absolute Difference (OSAD) between the bloc reference and the similar bloc.

Keywords: *Moving People Detection, Mog, FFT-Correlation, Optimized-SAD, Bounding Box,*

1. INTRODUCTION

The desire of safety applications in public environment has been increasing rapidly since September 11. Especially at sensible environments such as banks, airports, government buildings and in all other area where many people gather. For security purposes, this environment usually uses a fixed camera. This video surveillance system adopted should be monitored by human operators. Practically, monitoring large places requires many operators at the same time and along 24 hours, and it's not a simple task.

People tracking is the process of locating a moving people (or multiple peoples) over time using a camera. Tracking people in a video sequence is to determinate the position of the center of gravity and trace the trajectory, or to extract any other relevant information. There are a variety of algorithms used to perform video tracking, target representation and localization is mostly a bottom-up process. Locating and tracking the target object successfully is depend on the algorithm; for example, using blob tracking is used for identifying human motion etc.

The following common are some target representation and localization methods:

- Mean-shift tracking: an iterative localization procedure based on maximization of a similarity measure (Bhattacharyya coefficient).
- Contour tracking: detection of object boundary box (active contours …).
- Visual feature matching
- Blob tracking : segmentation of object interior (for example block based correlation)

Actually, many methods for tracking people in sensible area have been proposed, and can be divided into three categories: region based tracking [1][2], tracking with shape information [3][4] and adaptive background generation-subtraction[5][6]. Several systems have been proposed to track people with a stationary camera; Segen and Pingali [7] have developed a system for real time tracking of people in video sequence. In their algorithm, the input data of the system is recoded or live video captured by a fixed camera and the output consists

Journal of Theoretical and Applied Information Technology
10th February 2013. Vol. 48 No.1

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

of trajectory which gives the spatio-temporal coordinates of the moving person.

Yang and Waibel in [8] and Wang and Chang in [9] focus on tracking using a single camera and both use skin color for segmentation. M. A. Ali, S. Indupalli and B. Boufama [10] proposes a correlation based matching technique for feature-based tracking, and many other techniques using histogram etc.

Different researchers use different algorithms for tracking moving person, our design process of people tracking system has been described as the block given bellow.

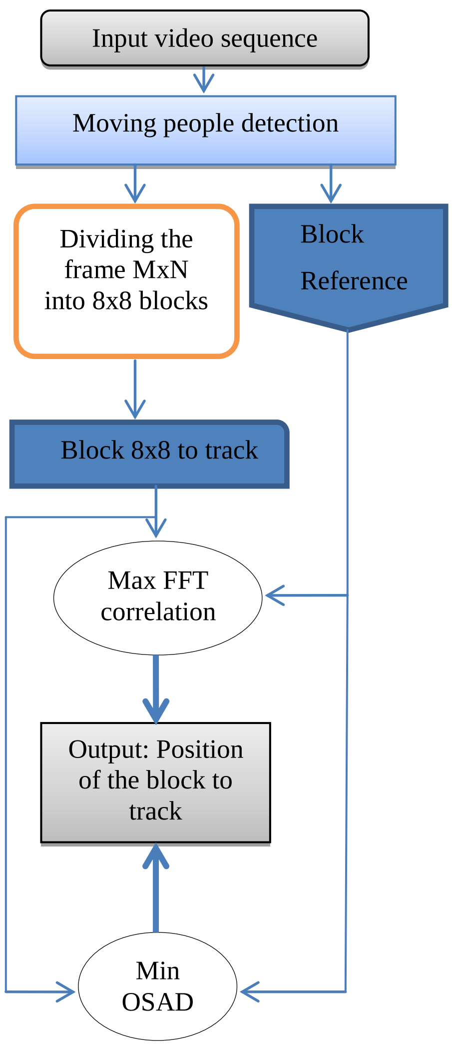


Figure. 1 The People Tracking Cycle

As seen from the above blocks, in this paper, we attempt to extract moving people with the Mixture of Gaussian algorithm, then we search the similar block to block reference, then we search the Max (Corr_FFT) and the Min (OSAD) for all blocks in the frame. The system return the position of our block tracked.

The remainder of this article is organized as follows. Section 2 gives a brief discussion in term of related work in order to explain the real issues of this study. Section 3 presents the process of moving people detection, we choose the Gaussian Mixture Model in order to do the process of detection, because it return a minimal misclassified pixels [15]. Section 3 gives the mathematical formulation of Bounding Box area of the people in motion; we surround the moving people by the bounding box in order to limit the search area in the Bloc Matching process. Fast Fourier Transform Correlation and the Optimized Sum of Absolute Difference formulas are used to determine the position of the searching bloc. The results of its application between blocs in Bounding Box of moving people are discussed in Section 4, in this section FFT-Correlation and OSAD curve are presented to explain the choice of the corresponding bloc tracked in the moving people matrix. The paper concludes and presents our perspective in Section 6.

2. RELATED WORK

There has been a considerable study activity on the people tracking systems from capturing video sequences over the last years. This section briefly reviews the process of people tracking in term of detecting the moving objects firstly and secondly tracking them with analyzing many types of significant information and related work to this project.

The people detection is the first important step in any surveillance system, in this process of motion segmentation; the pixels of each moving people are detected. Generally, the motion detection or motion segmentation consists of separating the background and foreground pixels. Frame differencing [16], medium filter [17] are simple and common techniques for background subtraction. Stauffer and Grimson [18] use the mixture of Gaussians to apply a two pass grouping technique to segment foreground pixels.

Following the process of background extraction, the moving people are tracked. Tracking people algorithm is very important in video surveillance domain and it can be categorized as either deterministic or stochastic based on their mechanism of searching. Stauffer and Grimson [18] use a linearly predictive multiple hypotheses to track moving objects. Yang et al [16] employ a

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

correspondence matrix and splitting algorithm to relate the measured foreground regions to the tracked objects. In [19], a technique for dynamically updating the models to accommodate changes in apparent color due to varying lighting conditions was used to track people in varying lighting conditions. In [20], many other techniques are proposed in video processing domains. Especially, for motion estimation, object tracking and segmentation. Tracking people in motion is a complex task and a variety of algorithms have been established to solve this problems. Until now, this task has not been definitively resolved and several workshops are open around. In this context, our work is born to help affectively to track people efficiently.

3. PEOPLE DETECTION

Among the high complexity methods, the Gaussian Mixture Model dominates the literature; it has their advantages and inconvenient. It returns a foreground with minimal errors in term of misclassified pixels, in this model the each pixel location is represented by a number (or mixture) of Gaussian functions that sum together to form a probability distribution.

In the MoG formulation [11], each pixel is modeled by K-Gaussians distributions, the probability of observing the current pixel value (pixel intensity variation) is:

$$P(XN) = \sum_{j=0}^{k} w_j\, \eta(XN, \Sigma_j) \qquad (1)$$

Where k is the number of distributions, each Gaussian is represented by their values (μ_k, σ_k, w_k) respectively, the mean vector, the covariance matrix and the weights.

Where η is a Gaussian probability density function:

$$\eta(X, \mu_k, \Sigma_k) = g_{(\mu,\Sigma)}(X) = \frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma_k|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1}(x-\mu_k)} \qquad (2)$$

μ_k, σ_k, and w_k are estimated by the Expectation Maximization algorithm (EM) [12] which maximizes the log-likelihood criterion.

4. MATHEMATICAL FORMULATION FOR TRACKED WINDOW

3.1 Bounding box for interest moving people

In this paragraph, we have to minimize the search window with surrounding the people in motion by a bounding box rectangle. The bounding box can be determined by computing the maximum value of x and y coordinates of interest moving object according to the following equation:

$$B^i_{min} = \left\{\left(x^i_{min}, y^i_{min} \right) | x,y \in O^i\right\} \quad (3)$$

$$B^i_{max} = \left\{\left(x^i_{max}, y^i_{max} \right) | x,y \in O^i\right\} \quad (4)$$

Where O^i denotes the set of the coordinate of points in the interest moving object i, B^i_{min} is the left-top corner coordinates of the interest moving object i and B^i_{max} is the right-bottom corner coordinates of the interest moving object i.

After surrounding the moving people in our sequence, we look for dividing the reduced matrix MxN from size (frame).

3.2 Correlation with Fast Fourier Transform

Correlation is the close mathematical cousin of convolution [13]. The model of FFT-Correlation applied in our work is written as in the equation:

$$\text{FFT\_Corr} = \sum_{i=0}^{N-1} X(i)Y(i+k) \qquad (5)$$

The correlation between a reference bloc frame and another in searching window form the interest moving object, it is maximal if we have a similarity between elements. We compute the correlation as follows: FFT the two data sets of blocs, multiply one resulting transform by the complex conjugate of the other, and inverse transform the product. The result will formally be a complex vector of length N. But, it will turn out to have all imaginary values zero since the original data. The values of the complex vector are the values of the correlation.

3.3 Optimized Sum of Absolute Differences

Recently, this algorithm is widely used; it's very easy and simple in order to find the similarity and each other relationship between the image blocs. It based on the corresponding point in the template bloc will be used for comparison. We can use this algorithm in many applications such as: object recognition, video compression, motion estimation and tracking. The following equation gives Optimize Sum of Absolute Differences (OSAD) [14]:

$$d(X,Y) = \sum_i \sum_j \frac{|X(i,j) - Y(i,j)|}{\max(X(i,j), Y(i,j))} \qquad (6)$$

Generally, the OSAD maximize the error rate.

Journal of Theoretical and Applied Information Technology
10th February 2013. Vol. 48 No.1

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

5. EXPERIMENTAL RESULTS

The proposed technique for tracking the template block in interest moving people was tested in video sequence extracted from video PETS2006 (Ninth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance) between different frames.

In this section, we present results of frames extracting from our PETS scene used and diagrams for Fast Fourier Transform and Optimized Sum of Absolute Differences.

-a- -b-

Figure.2 –A- Original Frame And –B- After Bounding Box Surrounding The People In Motion

During this step, we convert the original RGB frames to Gray Scale system with the following equation:

G_S_I[] = 0.229*R + 0.587*G + 0.114*B

In the last, we have to play with one matrix. Our tracking bloc is narrow at the bottom-right of view, and then only the window containing people is considered as in the following images:

-a- -b-

Figure. 3 -A- Extracting View And –B- Window For Tracked People

The bloc searched in the window of interest moving people is:

$$
Bref1 = \begin{matrix}
67 & 67 & 67 & 67 & 71 & 56 & 47 & 55 \\
73 & 74 & 76 & 78 & 78 & 63 & 52 & 58 \\
71 & 74 & 77 & 78 & 80 & 62 & 51 & 61 \\
69 & 73 & 75 & 77 & 74 & 58 & 48 & 60 \\
73 & 75 & 76 & 76 & 75 & 58 & 48 & 60 \\
75 & 75 & 75 & 74 & 83 & 65 & 53 & 62 \\
70 & 75 & 80 & 83 & 82 & 69 & 60 & 69 \\
70 & 73 & 77 & 78 & 86 & 71 & 60 & 69
\end{matrix}
$$

We have to calculate the Optimized error SAD; all values are saved in a vector, in parallel, we calculate the Correlation basing Fast Fourier Transform and save their values in one other vector. Finally, by extracting of the maximal of the OSAD and the minimal value of our vector OSAD of Corr_FFT as in the followings diagrams:

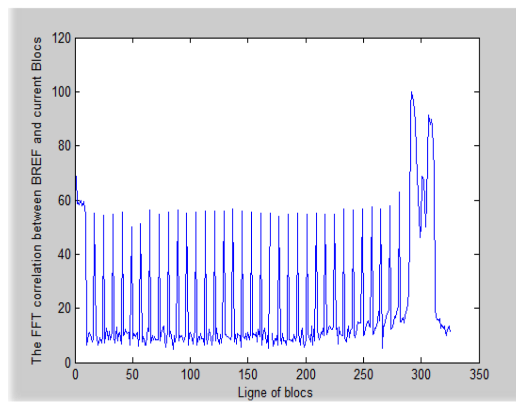


Figure .4 Correlation FFT Values

As we conclude in the above curve maximal values: Corr_FFT [bloc (292)] = 98.8218) are centered in the bloc number 292.

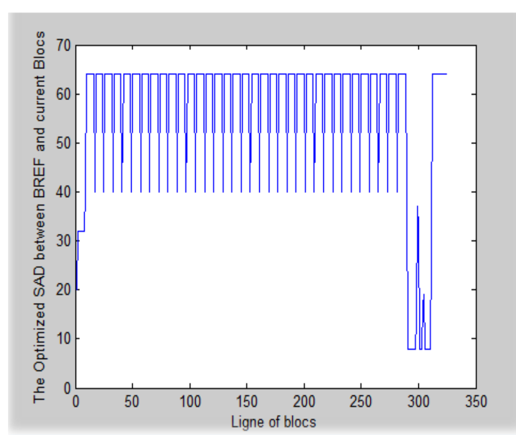
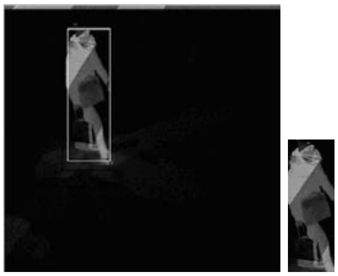


Figure .5 Optimized SAD Values

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

The minimal error value is in the same bloc 292.

Applying the Optimized SAD we return the same similar block to our block reference number 292 with minimal OSAD. The structure and position of the tracked block changes with displacement of our interest people in the scene.



-a- -b-

Figure .6 PETS2006 Video, -A- View Frame In Grayscale And –B- Extracted People In Motion In Boundary Box Rectangle

Bref2 =

149	150	143	134	124	98	61	36
147	150	147	142	135	113	78	52
144	149	148	146	144	132	102	78
145	148	146	145	150	144	120	99
145	146	143	144	148	140	117	98
143	143	139	141	141	126	99	79
147	143	140	140	138	115	84	68
155	149	143	145	138	113	83	70

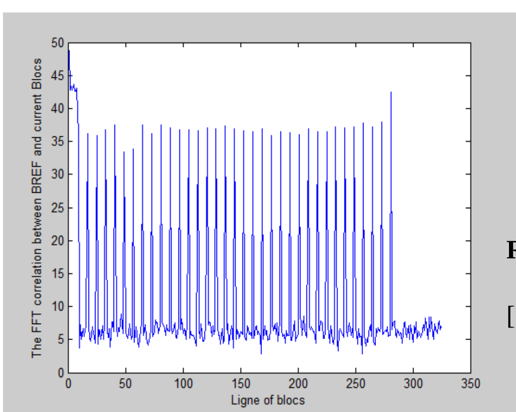


Figure .7 Correlation FFT values

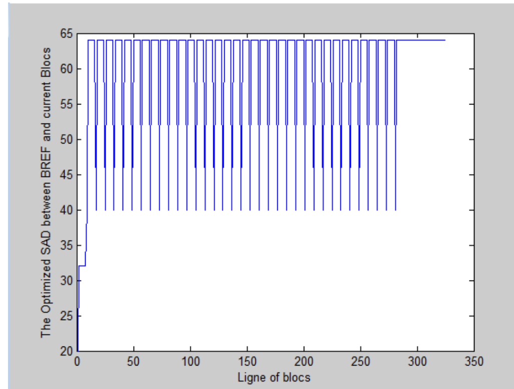


Figure .8 OSAD Errors

In this diagrams, we notes that the max(Corr_FFT) = 48.7904 and min(VOSAD) = 20 are obtained in block number 1.

CONCLUSION

In this paper, we have proposed a new technique which uses the Correlation based Fast Fourier Transform and the Optimized Sum of Absolute Difference error in order to track simply a person in a captured video sequence. Our method resume the time of procuring all frames in the captured video sequence into a short time of limited region of bounding box. This minimal bounding box is considered composed by the maximal of pixels of moving person.

Finally, we conclude from experimental results that all tracked bloc coincides with the minimal value of OSAD error and the peak in the big value of FFT-Correlation returned during the bloc matching process.

In the future studies we will attempt to generalize this technique to many peoples in motion with resolving the problem of occlusion in first time. Secondly, we will take different sizes of our tracked block, testing them in many video sequences.

REREFRENCES:

[1] G.D. Hager and P. N. Belhumeur,"Efficient Region Tracking with Parametric Models of Geometry and Illumination", IEEE Trans. On Pattern and Analysis Machine Intelligence, Vol. 20, N°. 10, pp. 1025-1039, October 1998

[2] P. Chang and J. Krumm, "Object Recognition With Color Cooccurrence Histograms," IEEE Conf. On Computer Vision and Pattern Recognition, Fort Collins, CO, June, 1999.

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

[3] A. Blake, M. Isard, "Active Contours", Springer, London, England, 1998.

[4] A. Koschan,S. K. Kang, J. K. Paik, B. R. Abidi and M. A. Abidi, "Video Object Tracking Based on Extended Active Shape Models with Color Information", Proc. 1st European Conf. Color in Graphics, Imaging, Vision, pp. 126-131, University of Poitiers, France, April, 2002.

[5] I. Haritaoglu, D. Harwood and L. S. Davis, W4 : Rea l-Time Surveillance of People and Their Activities," IEEE Trans. On Pattern Analysis and Machine Intelligence, Vol. 22, No. 8, pp. 809-830, August 2000.

[6] F. Dellaert and R. Collins, "Fast Image-Based Tracking by Selective Pixel Integration," ICCV 99 Workshops on Frame-Rate Vision, Sept, 1999.

[7] J. Segen and S. Pingali, "A Camera-Based System for Tracking People in Real Time," IEEE Proc. of Int. Conf. Pattern Recognition, Vol.3, pp.63-67, 1996.

[8] J. Yang and A. Waibel, "A Real Time Face Tracker, "IEEE Workshop Appl. Comput. Vision, pages 142–147, Los Alamitos, CA USA 1996

[9] H. Wang and S. Chang, " A Highly Efficient System for Automatic Face Region Detection in MPEG Video," IEEE Transactions on Circuits and Systems for Video Technology, 7(4):615–628, August 1997.

[10]M.A.Ali,S.Indupalli and B. Boufama,"Tracking Multiple People for Video Surveillance,"VP4-S06 The first international workshop on video processing for security, June 7-9 /2006, Quebec, Canada

[11] C. Stauffeer and W. E. L. Grimson : Adaptive Background Mixture Models for Real-Time Tracking in: Computer Vision and Pattern Recognition Fort Collins, Colorado (Jun.1999) pp. 246-252.

[12] A Dempster, N. Laird, and D. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm," Journal of the Royal Statistica l Society, 39 (Series B): 1-38, 1977

[13] Numerical Recipes in C: The art of Scientific Computing (ISBN 0-521-43108-5) Cambridge University Press Programs Copyrights ©1988-1992

[14] N. N. Dawoud, B. B. Samir and J. Janier," Fast Template Matching Method based Optimized sum of Absolute Difference Algorithm for face Localization. International Journal of Computer Applications (0975 – 8887) Volume 18-N°. 8, March 2011

[15] A. Ezzahout, R. Oulah Haj Thami and M. Y. Hadi,"Quantitative and Qaulitative evaluation based on misclassified pixels for people's detection" 2012 International Conference on Complex Systems, Agadir, Morocco.

[16] T. Yang, S. Z. Li, Q. Pan, and J. Li, "Real-time Multiple Objects Tracking with Occlusion Handling in Dynamic Scenes," in Proceeding of the 2005 IEEE Computer Society on Computer Vision and Pattern Recognition, vol. 1, 2005, pp. 970–975

[17] R. Cuccharina, C. Grana, M. Piccardi, and A. Prati, "Detecting Moving Objects, Ghosts, and Shadows in Video Streams, " in Transactions on Pattern Analysis and Machine Intellegence, 2003.

[18] C. Stauffer and W. E. L. Grimson, "Learning Patterns of Activity Using Real-Time Tracking, " in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, pp. 747–757.

[19[19] Y. Raja, "Tracking and segmenting people in varying lighting conditions using colour" Pages : 228 - 233, in Automatic Face and Gesture Recognition. 1998 Proceedings, Third IEEE International Conference 14-16 Apr 1998