Human Detection and Tracking in Video Surveillance System

Sriram K.V.
Assistant Professor, Department of E&C Engineering, Angadi Institute of Technology and Management, Belagavi/ Research Scholar, Visvesvaraya Technological University, Belagavi, Karnataka, India.
Email-shreeramkv@gmail.com

R. H. Havaldar
Professor,
KLE Dr. M.S.Sheshgiri College of Engineering and Technology, Belagavi,
Karnataka, India.

***Abstract*—The human detection and tracking in a video plays major roll in security systems. This paper proposes an approach to detect and track the persons in a video. This approach uses Gaussian Mixture Model to detect the person and Kalman filter to track the detected person. The processing time to detect the person is reduced by performing the detection operation on down-sampled video. After detecting the person, the original size of the video is reconstructed using Papoulis-Gerchberg method. The performance analysis is carried out by comparing with the state-of-the-art-algorithms. The experimental results show that the proposed method is well suited for detecting and tracking the person in lower processing time.**

Keywords*—Gaussian Mixture Model; Super-Resolution; Region of Interest; Kalman filter.

I. INTRODUCTION

In recent years, detecting moving people in a video picture is drawing attention of researchers because of large range of applications in people counting, people tracking, people identification etc. In such application areas, the background basically remains the same. So the moving persons are the important parts, which are defined as region of interest (ROI). The authors proposed [1] [2] a method to detect pedestrians using Haar-like features which are used in face detection method [3]. The author [1] proposed framework where he used four example-based detectors to localize the hands, head and legs on a human body. The author [4] [5] proposed color self-similarity (CSS) feature which learns the color structures of human. This feature is useful complement to the Histogram of Oriental Gradients descriptor (HOG) feature. Huge amount of video data in video surveillance make manual initialization impractical. Brendel [6] and Brox [7] proposed method to utilize an object's trajectory and motion cues. But, these methods do not have a precise model of the object and so the segments usually do not correspond to an actual object. Background subtraction [8-10] is an efficient way to detect moving objects. It extracts the foreground regions by calculating the difference between a video frame and the background model. The author [11] proposed, a method to detect the skin color of moving person present in a static background, and using non-Gaussian recursive Bayesian Particle filter person tracking is carried out. The author [12] proposed a method for detection and tracking of moving person. This approach is tested on Weizmann dataset containing run, sleep, bend, etc. activities. The author [13] proposed an approach to detect and track the human mouth region based on the color components of skin and lips.

In this proposed paper, the area of moving person is detected by Gaussian Mixture Model (GMM) [14] [15] and regarded as ROIs to be processed further. This work also uses super resolution technique to reduce the computation time for detection of person in a scene.

Section II describes proposed person detection method. Section III presents the results obtained using proposed method and comparison to those produced by other methods. The paper is concluded in Section IV.

II. PROPOSED METHOD

The proposed approach is as shown in Fig.1. The proposed method is as discussed below.

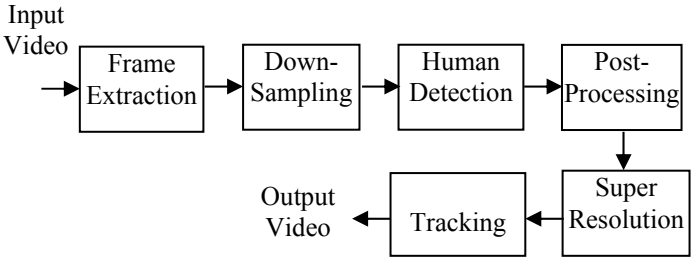


Fig.1. Proposed person detection and tracking system

Frame Extraction and Down-sampling:

RGB frames are extracted from a video. These frames are down-sampled by factor 2. This down-sampling method reduces the execution time required to detect a moving person in a frame.

Person Detection:

Background subtraction algorithm is applied to detect Foreground object and Background pixel. The Gaussian Mixture Model [15] is applied in this proposed method. Here, every pixel is modeled as a mixture of k Gaussian ($k = 3$).

$$P(x) = \sum_{i=1}^{l} w_{(i)} G(x, \mu, t, z) \quad (1)$$

978-1-5090-0612-0/16/$31.00 ©2016 IEEE

$where\ l$ - Gaussian clusters number, $w_{(i)}$ - Weight factor correlated with cluster. G - Gaussian pdf, μ –mean and z- covariance matrix of Gaussian cluster

Post processing

The result of above method may consist of incomplete shape, noise, etc. Such problems are solved by using morphological filter. The detected object is characterized by its centroid and the regular shape around the object boundary.

Super-resolution

The processed frames are enhanced using super-resolution technique to get its original size. Super-resolution (SR) is a technique which reconstructs high resolution (HR) image using set of low resolution (LR) images [16]. The observation model of SR is

$$Y_k = DB_kM_kX + N_k \quad (2)$$

where, $k = 1, 2, 3, \ldots K$` . D – Factor of down-sampling, M_k - motion vectors B_k - blur matrix and Nk – additive noise.

The frames are super-resolved using Papoulis-Gerchberg [17][18]method. The modified PG algorithm is as below,

i. Divide frame in $n \times n$ patches
ii. Select the patch
iii. Create a HR trellis
iv. The reference frame of first LR image is set as pixel place and then put in the identified pixels of LR images.
v. Transform patch into frequency domain
vi. In the frequency domain, the high-frequency components are set to zero
vii. Transform back to spatial domain. For all the patches in the frame, perform above mentioned steps.

Tracking

The motion of the detected object is tracked by using Kalman filter [19]. This filter predicts the location of object in each frame. This method reduces the computation time.

Quality Measures

The proposed algorithm is quantitatively evaluated by using similarity measure (SM) [20] [21].

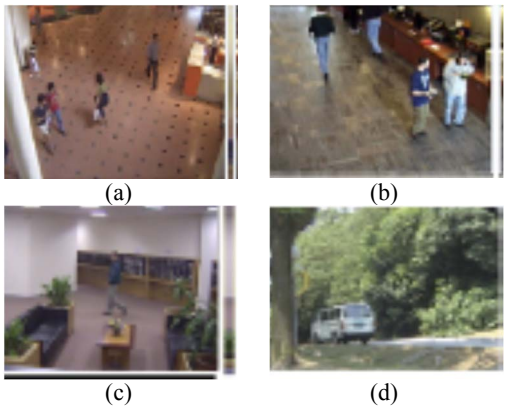
The similarity measure is defined as follows:

$$Q(x, y) = \frac{n(x \cap y)}{n(x \cup y)} \quad (3)$$

where x-foreground region, y-ground truth, n- number of pixels

III. RESULTS

The proposed work is implemented using Matlab2009a on Intel Core i3 machine with processor speed of 1.8GHz and RAM size of 4 GB. The performance of the proposed approach is evaluated by using four benchmark video sequences Shopping center, Buffet Restaurant, Lobby & Campus. The input video sequences are as shown in Fig.2. From a video, fifteen frames were randomly selected for testing purpose.



(a) (b) (c) (d)

Fig.2. Input video sequences (a) Shopping center (b) Buffet Restaurant (c) Lobby (d) Campus

The proposed approach is compared with algorithms proposed by Huang [22] and Tsai [23]. The averaging values of similarity measures for benchmark video sequences are shown in Table I

Table I. Analysis of proposed method

S. No.	Video	Huang		Tsai		Proposed	
		SM	*Time (Sec.)*	*SM*	*Time (Sec.)*	*SM*	*Time (Sec.)*
1	Campus	0.59	9	0.63	8	0.83	5
2	Shopping center	0.67	13	0.75	10	0.88	7
3	Buffet Restaurant	0.52	8	0.61	8	0.77	6
4	Lobby	0.49	7	0.55	5	0.71	6

From the Table I, it can be seen that, similarity measure of proposed method is good compared to other methods and this method detects human in very less time compared to state-of-the-art-algorithms.

IV. CONCLUSION

The goal of this work is to detect and track human in a video sequence. In this paper, a method is proposed to detect and track the persons in a video. This method consists Gaussian Mixture model to detect the human and Kalman filter is used to track the human in video sequence. The computation time is reduced by using super-resolution technique. From the analysis, it can be seen that the proposed method produces better results compared to state-of-the-art-algorithms. The detection results and minimalism of algorithm make the proposed approach to be an appropriate method for real-time application.

REFERENCES

[1] A. Mohan, C. Papageorgiou, and T. Poggio, "Example-based object detection in images by components," IEEE Transactions on Pattern Analysis and Machine Intelligence, volume 23, 2001

[2] P. Viola, M. J. Jones, and D. Snow, " Detecting pedestrians using patterns of motion and appearance," In ICCV, 2003.

[3] P. A. Viola and M. J. Jones, "Robust real-time face detection," International Journal of Computer Vision, volume 57, 2004.

[4] S. Walk, N. Majer, K. Schindler, and B. Schiele, "New features and insights for pedestrian detection," IEEE Conference on Computer Vision and Pattern Recognition, June 2010.

[5] Woonhyun Nam. "Improving object localization using macrofeature layout selection", 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), 11/2011

[6] W. Brendel and S. Todorovic, "Video object segmentation by tracking regions," IEEE International Conference on Computer Vision, 2009.

[7] T. Brox and J. Malik. "Object segmentation by long term analysis of point trajectories," Springer Conference of computer vision, 2010.

[8] H. Sidenbladh and M. J. Black, "Learning the statistics of people in images and video," International Journal of Computer Vision, 2003.

[9] C. Stauffer and W. Grimson. Adaptive background mixture models for real-time tracking. In CVPR, 1999.

[10] K. Toyama, B. Krumm, J.and Brumitt, and B. Meyers. Wallpower, "Principles and practice of background maintenance," In ICCV, 1999.

[11] Sharma, Aashish, Ajay Singh, and Rajesh Rohilla. "Color based human detection and tracking algorithm using a non-Gaussian adaptive Particle filter", 2016 3rd International Conference on Recent Advances in Information Technology (RAIT), 2016.

[12] Rajvir Kaur, Sonit Singh "Background modelling, detection and tracking of human in video surveillance system " Innovative Applications of Computational Intelligence on Power, Energy and Controls with their impact on Humanity, 2014.

[13] C. Anitha, M.K. Venkatesha, B. Suryanarayana Adiga. "Real Time Detection and Tracking of Mouth Region of Single Human Face", 2015 3rd International Conference on Artificial Intelligence, Modelling and Simulation (AIMS), 2015.

[14] Maghsoumi, Hossein, Davud Asemani, and Hadi Amirpour. "An efficient adaptive algorithm for motion detection", 2015 IEEE International Conference on Industrial Technology (ICIT), 2015.

[15] C. Stauffer, W.E.L. Grimson, "Adaptive Background Mixture Models for Real-Time Tracking," Proc. Computer Vision and Pattern Recognition Conf., 1999

[16] Anand Deshpande and Prashant Patavardhan, "Multiframe Super-Resolution For Long Range Captured Iris Polar Image," IET Biometrics, 2016. DOI-10.1049/iet-bmt.2016.

[17] Priyam Chatterjee and Sujata Mukherjee, "Application of Papoulis-Gerchberg Method in Image Super-resolution and Inpainting," The Computer Journal, Vol.1, 2007.

[18] Deshpande, Anand, Prashant P. Patavardhan, and D. H. Rao. "Super-resolution for iris feature extraction", 2014 IEEE International Conference on Computational Intelligence and Computing Research, 2014

[19] Garry A. Einicke and Langford B. White, "Robust Extended Kalman Filtering," IEEE Transactions On Signal Processing, Vol. 47, No. 9, September 1999

[20] L Li, W Huang, IY Gu, Q Tian, "Statistical modelling of complex backgrounds for foreground object detection," IEEE Trans. Image Process. 2004.

[21] Kermani and Asemani "A robust adaptive algorithm of moving object detection for video surveillance," EURASIP Journal on Image and Video Processing 2014, 2014

[22] Li, L., Huang, W., Gu, I., and Tian, Q. "Statistical modeling of complex backgrounds for foreground object detection," IEEE Transactions on Image Processing, 2004.

[23] TH Tsai, CY Lin, "Markov random field based background subtraction method for foreground detection under moving background scene," Conference on Genetic and Evolutionary Computing, 2010.