An Approach to Deformable Object Tracking in Video

Chandra Mani Sharma
Department of IT
ITS Mohan Nagar
Ghaziabad

Harish Kumar, Ph.D
Department of IT
ITS Mohan Nagar
Ghaziabad

Deepika Sharma
Department of IT
GBS Patparganj
New Delhi

ABSTRACT
Object tracking refers to spotting the presence of one or more objects of interest in each frame of a video. The task of object tracking is mainly useful in visual surveillance and scene understanding applications. With the increasing availability of rich video contents, new application areas are emerging every day. Generally, object tracking plays an intermediate role in many such applications. Several techiques can be found in literature but the basic evaluation and choice parameter that many applications consider include fast execution speed and automatic operation of the tracking method. There remains a tradeoff between accuracy and execution speed for the object tracking methods. Further it is true that in today's applications the most important category of objects is human himself. This paper proposes a technique for human detection and tracking in video. The proposed method is accurate and efficient in execution speed. Several experimental results presented in the paper demonstrate novelty of method.

General Terms
Pattern Recognition, Security, Algorithms.

Keywords
Human Detection, Object Tracking, Visual Sueveillance, Video Processing

1. INTRODUCTION
In recent past among all types of data available, the video data has surpassed all in quantitiy. Now it poses new challenges before the human civilization to meticuloisly anlyze and to organize the available memonth size of video data. The sources of this data include; the video stream captured by CCTV cameras, social networking sites, movies, clips captured by personal gadgets etc. Tracking the objects of interest in a video sequence plays the role of a good starting point for innumerable applications. There could be traced various applications of video scene analysys [1-4].

Hu et al. [1], surveyed on visual surveillance and its various applications. The problem of multiple object tracking is more challenging than single object tracking [5]. It is because of sudden appearance and disappearance of object in a video frame. It is better to detect objects before tracking them for enhanced accuracy [6]. In [7] a technique called adaptive boosting is used to speed up a binary classifier traiing [7]. Our experimental results demonstrate that *particle filtering* is useful for object tracking because of its capability to effeciently represent object tracking system (of non-linear type) in the presence of non-Gaussian noise. Here we refer some of the object tracking techniques found in literature [8-12]. These approaches fall under two main categories; single object trackers and multiple object trackers. Lanvin et al. [8], proposed an object detection and tracking technique and solved the non-linear state equations using particle filtering. Algorithms which attempt to find the target of interest without using segmentation have been proposed for single target tracking based on cues such as color, edges and textures [13].

However single object trackers suffer from the problem of false positives, when severe occlusion occurs because of hidden first order Markov hypotheses [9]. The problem of tracking multiple objects using particle filters can be solved in two ways. One is by creating multiple particle filters for each track and another one is by having a single particle filter for all tracks. The second approach works fine as long as the objects under tracking are not occluded but in case of occlusion when objects come close by, these techniques fail to track the objects. In [9] a multi-object tracking technique using multiple particles has been proposed. Chen et al. [10] proposed a color based particle filter for object tracking. Their technique is based on Markov Chain Monte Carlo (MCMC) particle filter and object color distribution. In [11], object tracking and classifications are performed simultaneously. Many of the particle filters based multiple object tracking schemes rely on hybrid sequential state estimation. The particle filter developed in [14] has multiple models for the objects motion, and comprises of an additional discrete state component to denote which motion model is active. The Bayesian multiple-blob tracker [15] presents a multiple tracking system based on statistical appearance model. Multiple blob tracking is managed by incorporating the number of objects present in the state vector and the state vector is augmented as in [16] when a new object appears in the view.

The rest of the paper is organized as follows. Section2 discusses the proposed technique for human detection and tracking, in section3 the various experimental results using the proposed technique are given which prove the validity and novelty of the method, section4 comprises of the conclusion.

2. THE PROPOSED METHOD FOR TRACKING
The detection and tracking of deformable objects (objects which can change their shapes and are non-rigid) in video is far more challenging than detection and tracking of rigid objects. Here we solve two problems. One is to detect the humans in video and another one is to track them in the subsequent video frames. The sub problem of object detection is solved using machine learning approach for which we train our human detector using binary adaptive boosting and for

tracking a particle filter is use. The basic steps of the method are given below:

- Collect and preprocess the image samples.
- Perform binary adaboost training using the Haar-like features extracted from the training samples.
- Create particle filter for tracking.
- Create motion and appearance models.
- Human detection and tracking in video frames.

The proposed algorithm is given below:

Step 1: Let Z be the input video to the algorithm.

a) In first frame η_0 of Z, detect humans using the human detector. Let ω be the number of detected humans.

b) Initialize trajectories δ_j, $1 \leq j \leq \omega$ with initial positions $\alpha_{j,0}$ of the human beings detected by the detector and also set the occlusion count ϕ_j for each of these trajectories to 0.

c) Initialize the appearance model ψ_j for each trajectory from the region around $\alpha_{j,0}$.

Step 2: For each subsequent frame η_i of input vide Z,

For each existing trajectory δ_j,

i. Use motion model to predict the distribution $\theta(\alpha_{j,i} | \alpha_{j,i-1})$, over locations for human j in frame i, creating a set of candidate particles $\alpha_{j,i}^{(k)}$, $1 \leq k \leq K$.

ii. Use appearance model to compute the color histogram $\psi^{(k)}$ and likelihood $\theta(\psi^k | \alpha_{j,i}^{(k)}, \psi_j)$ for each particle k.

iii. Acquire k^*, the index of the most likely particle, after re-sampling the particles according to their likelihoods.

iv. Now run the human detector on the location $\alpha_{j,i}^{(k^*)}$. If the location is classified as a human, reset $\phi_j \leftarrow 0$ else increase $\phi_j \leftarrow \phi_j + 1$

v. If ϕ_j exceeds a threshold, remove trajectory j.

Step 3: Now for frame η_i, search the new human objects and compute the Euclidean distance $\Omega_{j,k}$ between each newly detected human k and each existing trajectory δ_j. When $\Omega_{j,k} > \tau$ for all j, initialize a new trajectory for detection k. where τ is a threshold in pixels whose value is less than the width of the tracking window.

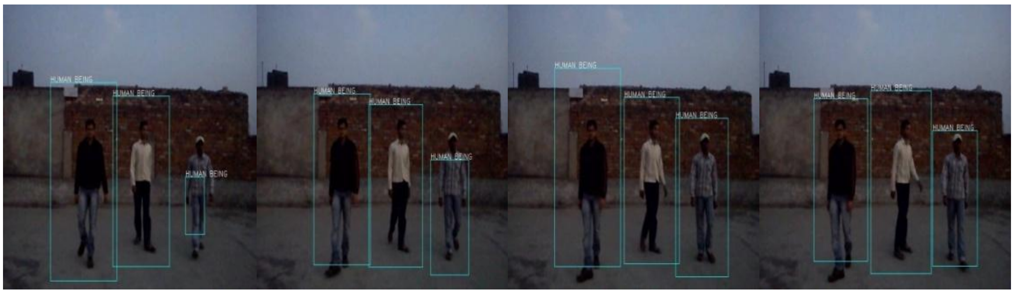
3. EXPEIMENTAL RESULTS

We have tested the proposed automatic human detection and tracking technique on a number of realistic videos. The human detection and tracking results with some of the representative videos are given in fig. 1 and fig. 3. Human detection and tracking using the proposed method starts automatically without providing any initialization parameters, unlike many other existing object tracking techniques in which operator intervention is required [5,17]. In fig.1 we have shown the human detection and tracking results with a realistic video consisting of three human objects walking towards the camera. The video was shot in evening time in dim lighting conditions at a resolution of 640×480 and a frame rate of 17 frames per second. The video consisted of total 500 frames. In fig. 1 the detection and tracking results for 16 frames, from frame 1to 375, at a difference of 25 frames are given. One can observe from fig. 1 that the proposed method detects and tracks all the three human objects from one frame to another with accuracy.

Frame 1 Frame 25 Frame 50 Frame 75

Frame 100 Frame 125 Frame 150 Frame 175

Frame 200 Frame 225 Frame 250 Frame 275



Frame 300 Frame 325 Frame 350 Frame 375

Fig. 1: Detection and tracking results with the proposed method on a real-time video in outdoor environment.

Table 1: Calculated and Actual Centroids of Human Objects during Tracking

	Calculated Object Centroids			Ground Truth Values for Object Centroids			Centroid Distance		
Frame No.	Object1	Object2	Object3	Object1	Object2	Object3	Object1	Object2	Object3
1	(473,307)	(368,297)	(201,283)	(473,307)	(368,295)	(213,290)	0.00	2.00	13.89
25	(471,304)	(367,297)	(200,282)	(461,307)	(364,295)	(212,289)	10.44	3.60	13.89
50	(479,316)	(373,306)	(214,317)	(479,316)	(373,306)	(214,317)	0.00	0.00	0.00
75	(489,310)	(370,314)	(209,308)	(478,305)	(372,319)	(210,307)	12.00	5.38	2.23
100	(481,310)	(381,316)	(218,309)	(479,311)	(381,316)	(218,309)	2.20	0.00	0.00
125	(490,322)	(388,306)	(215,310)	(494,320)	(384,306)	(215,310)	4.47	4.00	0.00
150	(486,340)	(384,312)	(227,300)	(484,340)	(385,311)	(237,309)	2.00	1.41	13.45
175	(489,314)	(372,313)	(232,304)	(487,311)	(375,314)	(231,309)	3.60	3.16	5.09
200	(489,341)	(376,315)	(227,320)	(488,344)	(375,314)	(224,321)	3.16	1.41	3.16
225	(477,312)	(388,334)	(217,303)	(473,310)	(387,335)	(215,304)	4.47	1.41	2.23
250	(478,313)	(390,337)	(220,300)	(478,313)	(390,337)	(220,300)	0.00	0.00	0.00
275	(471,304)	(390,337)	(220,300)	(468,304)	(390,337)	(220,300)	3.00	0.00	0.00
300	(473,307)	(368,334)	(201,283)	(475,308)	(368,334)	(200,281)	2.23	0.00	2.23
325	(470,302)	(368,334)	(220,300)	(468,300)	(368,334)	(219,298)	2.82	0.00	2.23
350	(470,302)	(368,334)	(220,300)	(468,296)	(368,334)	(220,300)	6.32	0.00	0.00
375	(466,298)	(367,332)	(220,300)	(462,290)	(366,331)	(220,300)	8.94	1.41	0.00

Table 1 gives the ground truth and calculated values of centroids of the bounding boxes, around the three objects in the video results shown fig 1. One can observe from the given Table 1 that the maximum centroid distance between the ground truth value and the calculated value is nearly 14 pixels (13.89) for the video of frame resolution 640×480, which is not too much with respect to the frame size. Most of the times the centroid distance is either 0 or is too negligible to cause any big deviation from ground truth values, showing algorithm's accuracy which is also evident from the centroid distance shown in fig. 2.

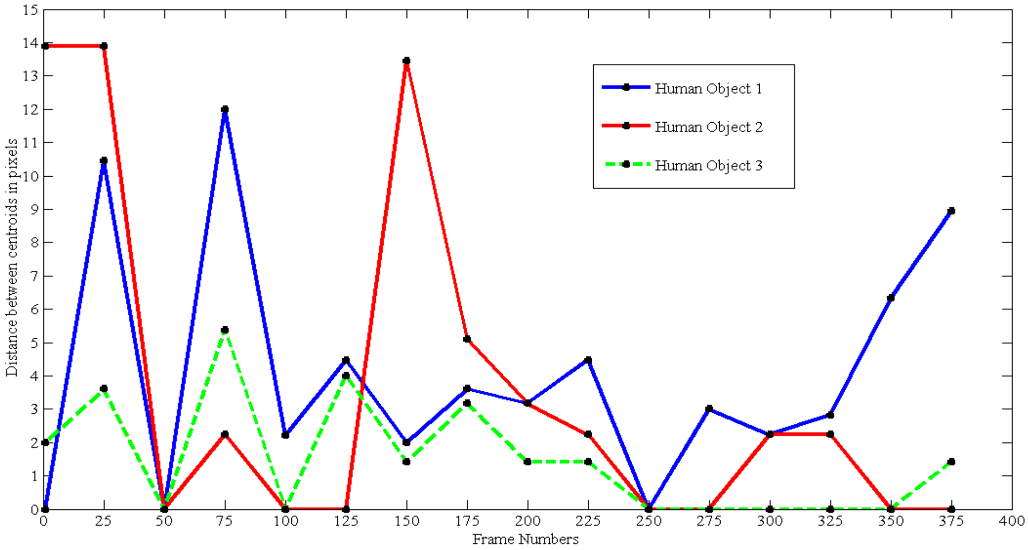


Fig 2: Grapg plot for tracking results shown in Fig. 1

Fig. 3: Human detection and tracking results with the proposed method when humans appear in different poses.

Fig. 3 shows the human detection and tracking results where the human may fall in different poses. There are four human beings in the video frame sequence in different poses. The four humans are in four different poses. The first one is walking, second one is standing , Third human is sitting and the fourth one is bending. Results show that our proposed algorithm can detect and track all four human beings even in this complex video.

The average detection and tracking accuracy of the proposed technique is 92.44%. The average execution speed is 18 frames per second. The method is quite suitable for outdoor environment and to detect & track deformable objects like human beings.

4. CONCLUSION

With the increasing number of crimes and terrorist activities the need for developing viable visual surveillance systems is being felt world wide. Object detection and tracking is the fundamental step in visual surveillance. In this paper we have proposed an automatic multiple human detection and tracking technique using machine learning and a simple particle filter. The operation of the proposed technique is fully automatic and does not require any operator intervention unlike other methods. Experimental results demonstrate that the technique can detect and track the humans in outdoor environment in poor lighting conditions and in multiple poses. The average detection accuracy 92.44% and the average processing speed of 18 frame/second, are the factors which make this technique robust for real-time visual surveillance application.

5. REFERENCES

[1] W. Hu and T. Tan, "A Survey on Visual Surveillance of Object Motion and Behaviors ", IEEE Trans. Systems, Man, and Cybernetics, Vol. 34, No. 3, pp. 334-352,2006.

[2] E.Stringa and C.S. Regazzoni, "Real-time Video Shot Detection for Scene Surveillance Application," IEEE Trans. Image Processing, Vol. 9, pp. 69-79, 2009.

[3] I. Haritaoglu, D. Harwood, and L. Davis, "W^4: Who? When? Where? What? a Real-time System for Detecting and Tracking People," In the Proc. IEEE Conference on Face Gesture Recognition, pp. 222–227, 1998.

[4] D. Koller, J. Weber, T. Huang, J. Malik, B. Rao G. Ogasawara, and S. Russell, "Toward Robust Automatic Traffic Scene Analysis in Real-time," In the Proc. IEEE Conference on Pattern Recognition, Vol. 1, pp. 126–131, 1994,.

[5] S. Nigam and A. Khare, "Curvelet Transform Based Object Tracking", In the proc. IEEE Conference on Computer and Communication Technology, pp. 230-235, 2010

[6] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," In the Proc. of IEEE Int. Conference on Computer vision and Pattern Recognition pp. 511-518, 2001.

[7] Y.Freund, and R. Schapire, " A decision theoretic generalization of on-line learning and an application to boosting," Journal of Computer and System Sciences 55, pp.119–139, 2007

[8] P. Lanvin, J-C.Noyer, M. Benjelloun," Object Detection and Tracking using the Particle Filtering," In the Proc. of IEEE 37th Asilomar Conference on Signals, Systems and Computers, Vol. 2,pp.1595-1599, 2003.

[9] J. Wang, Y. Ma, C. Li, H. Wang, J. Liu, "An Efficient Multi-Object Tracking Method using Multiple Particle Filters" In the Proc. of World Congress on Computer Science and Information Engineering, pp. 568-572, 2009.

[10] Y. Chen, S. Yu, J. Fan, W. Chen, H. Li, "An Improved Color-based Particle Filter for Object Tracking," In the Proc. of IEEE Second Int. Conference on Genetic and Evolutionary Computing, pp. 360-363, 2008.

[11] F. Bardet, T. Chateau, D.Ramadasan, "Illumination Aware MCMC Particle Filter for Long-term Outdoor Multi-Object Simultaneous Tracking and Classification," In the Proc. of IEEE 12th Int. Conference on Computer Vision, pp. 1623-1630, 2009.

[12] C. Luo, X. Cai, J. Zhang, "Robust Object Tracking using the Particle Filtering and Level Set Methods: A Comparative Experiment," In the Proc. of IEEE 10th Workshop on Multimedia Signal Processing, pp. 359-364, 2008.

[13] P. Brasnett, L. Mihaylova, N. Canagarajah, and D. Bull, "ParticleFiltering for Multiple Cues for Object Tracking in Video Sequences," In the Proc. of the 17th SPIE Annual Symposium on Electronic Imaging, Science and Technology, vol. 5685, pp. 430–440,2005

[14] M. Isard and A. Blake, "A Mixed-State Condensation Tracker with AutmaticModel-Switching," in Proc. Int. Conf. Computer Vision, pp. 107–112, 1998

[15] M. Isard and J. MacCormick, "Bramble: a Bayesian Multiple Blob Tracker," in Proc. Int. Conf. Computer Vision, pp. 34–41, 2001

[16] J. Czyz, B. Ristic, and B. Macq, "A Color-Based Particle Filter for Joint Detection and Tracking of Multiple Objects," In the Proc. of the ICASSP, pp. 217-220, 2005.

[17] T.B. Nguyen and A. Khare, "Object Tracking Of Video Sequence in Curvelet Domain", Int. Journal of Image and Graphics, pp. 1-20, 2011.