Precise Human Motion Detection and Tracking in Dynamic Environment

[1]Moiz A. Hussain,
ENTC Deptt.
Anuradha Engineering College,
Chikhli, India
mymoiz2004@yahoo.co.in

[2] G. U. Kharat,
Sharadchandra Pawar College of Engineering,
Otur, India

gukharat@gmail.com

Abstract: - An extensive amount of research is undergoing in the field of computer vision is related to object detection & tracking. Detection and tracking of the moving objects in dynamic environment is a difficult task. In this paper, we describe a framework for the detection and tracking of moving people in indoor\outdoor environment. The algorithm consists of two steps: at first, we use a Dual frame differencing method for motion detection. Then, algorithm performance is improved for false motion detection using morphological operations. The second step of the algorithm is object tracking framework based on Kalman filtering which works well in dynamic scenes. Experimental result shows that the method yields superior performance as compared to the other methods of video surveillance.

Key-words: - Video surveillance; Object Tracking; Kalman Filter; Dual Frame Differences.

1 Introduction

Object detection in videos is an important task for many higher-level visions processing such as motion capture, surveillance systems, and computer vision. Object tracking [10] is a difficult and challenging task that establishes the correspondence between object and their position in different frames of video. Tracking has been a tough task in the presence of background clutter and varying illumination condition. The objective of this work is to construct a system that manages to detect & track human motion in indoor as well as outdoor environment. Visual surveillance in dynamic scenes has a wide range of potential applications, such as a security guard for communities and important buildings, traffic surveillance in cities and expressways, detection of military targets, etc. A typical surveillance application consists of three building blocks: movement detection, object tracking and higher level motion analysis. The aim of the surveillance applications is to detect, track and classify targets.

The existing methods of Visual surveillance for object detection & tracking can be divided into three major categories: contour based models [1], region based models [2, 3], and feature point-based models [4, 5]. Also most of the programmed surveillance system uses frame differencing and background subtraction for object detection. This works on the principle that if there is change in pixel that is foreground & stationary pixels are background. But sometimes nonstationary background objects are misclassified as foreground. An intelligent automated tracking system should be able to detect changes caused by a new object, whereas non-stationary background regions, such as branches and leaves of a tree or a flag waving in the wind, should be identified as a part of the background.

Despite the fact that many different approaches have been proposed in the last decades, accurate objects tracking in dynamic environment is still one of the most challenging issues in the computer vision. There are a lot of difficulties for a single object tracking like illumination variability, background noise and occlusions. Multiple object tracking is even more challenging due to multi object occlusions. Changes in background illumination are a significant question of background extraction. Overcoming these background variations is an important issue in the accurate and fast extraction of the initial background.

In the simple subtraction method, the intensities of each pixel at adjacent points in time are subtracted from each other; nonzero values in the resulting difference image indicate that something in the image has changed. It is assumed that these changes are due to motion rather than other effect, such as illumination effect. However, the illumination can easily change that effect the image sequence even the camera is fixed in position. The illumination can change in an indoor and an outdoor image sequence. The illumination in an indoor sequence can change since the light changes at least 50 or 60 times per second due to indoor lighting; on the contrary, the image capture by a camera is 25 ~ 30 per second. Therefore, illumination has an effect on the captured image sequence. The illumination in an outdoor sequence also vary for different reasons such as the sun light change and this changes is very rapid than indoor sequence. The weather condition is another reason for changing the illumination.

The main contribution of this paper focuses on the method for tracking moving objects in an indoor & outdoor environment. We used dual frame difference method [9] to extract moving objects from a video sequence. However, this method suffers from the problem of false motion detection in dynamic background & hence we used this method in combination of some morphological operation & filters to achieve robust and accurate extraction of moving objects. How to design a simple and feasible system for detecting and tracking human motion and make it have a good robustness and real-time property is one of the research topics in the field of computer vision at present.

The technique presented in this paper is a solution to the problem of effectively detection and tracking of moving people with simplified model and less computational cost. One of the advantages of our system, compared to the other state-of-the-art methods is that it reduces the number of false detections, as pixel-level differentiation can be performed in regions with significant motion only.

The flow diagram of proposed system is shown in Figure 1.

The remainder of this paper is organized as follows: In Section 2 Database collection is presented. Section 3 describes the algorithm for moving target detection in an indoor & outdoor environment. Section 4 shows experimental results to demonstrate the accuracy of tracking in several datasets. Section 5 provides our future work. Finally, Section 5 concludes this paper.

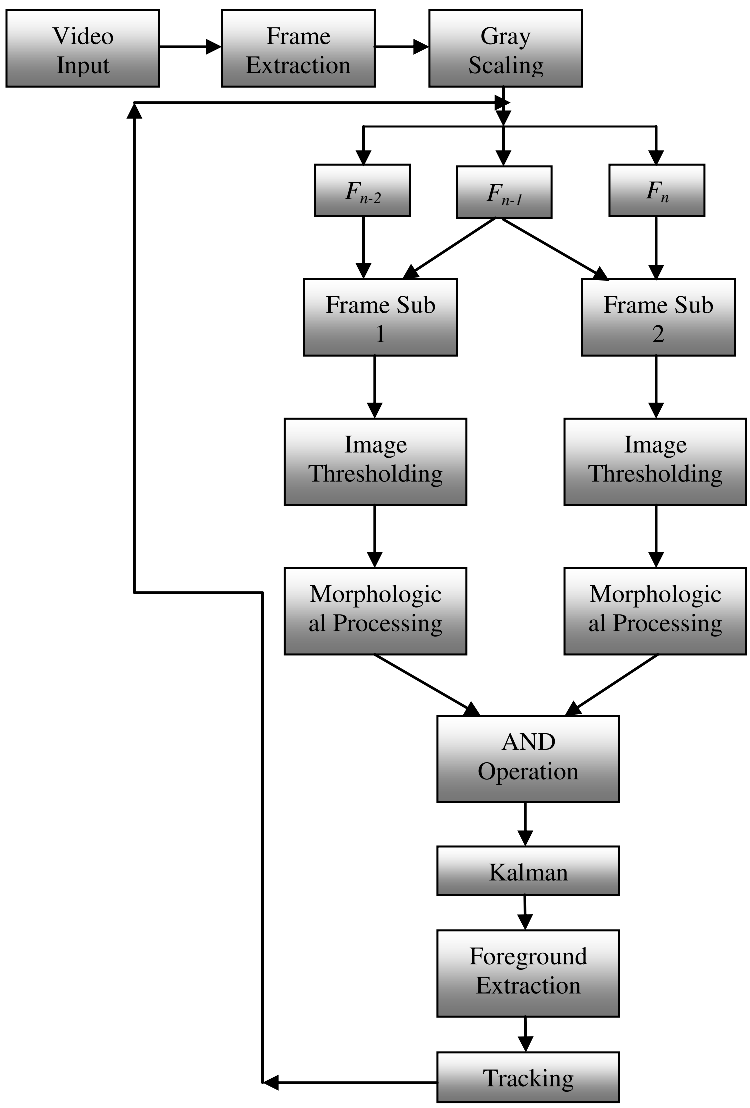


Fig.1: System Overview

2 Related Work

There is a rich literature on video surveillance for human motion detection & tracking.

Ross et al. [32] propose an adaptive tracking method that shows robustness to large changes in pose, scale, and illumination by utilizing incremental principal component analysis. The online multiple instance learning algorithms [33] successfully tracks an object in real time where lighting conditions change and the object is occluded by others.

The recent development of sparse representation [34] has attracted considerable interest in object tracking [35, 36] due to its robustness to occlusion and image noise. As these methods exploit only generative representations of target objects and do not take the background into account, they are less effective for tracking in cluttered environments. In

[37], the methods that are used in human action recognition were classified into global and local representation.

Avidan [41] trains a Support Vector Machine (SVM) classifier offline and extends it within the optical flow framework for object tracking. Collins et al. [42] use variance ratio of foreground and background classes to determine discriminative features for object tracking.

Spatial correction techniques [43, 44] produce balanced images that appear uniformly illuminated by a distant, diffuse source by removing spatially-varying illumination artifacts, such as shadows, highlights, and specular reflections. Because this approach focuses on spatial, intra-image lighting variations rather than temporal, inter image lighting changes, it is a different class of illumination compensation.

A mixture model of three components with an online EM algorithm is proposed to model the appearance variation during tracking [45]. The graph-cut algorithm was also used for tracking by computing an illumination invariant optical flow field [46].

The work by Cwojek et. al [50] proposed an approach for multiple human activity recognition in office environment with simultaneous tracking. In this research, both video and audio features are implemented to employ a multilevel Hidden Markov Model (HMM) framework. Pfinder [53] system was developed to describe a moving person in an indoor environment. It tracks a single non-occluded person in complex scenes. A Bayesian approach is proposed in the work of Li et al. [71]. Spectral features (color information), spatial features (statistics on pixel regions), and temporal features (filters to model temporal change of pixels) are incorporated in a Bayesian framework to characterize the background appearance at each pixel.

A graph illustration was recommended in [64] to represent the moving regions which were extracted from a video acquired by a moving airborne platform. This approach allows dynamic inference of moving objects to attain robust tracking and offers a confidence measure characterizing the reliability of each trajectory. The silhouette-based approaches utilize silhouettes in human movement analysis [65], as this is generally robust to image-noise disturbances. Viola [68] used spatio-temporal filters based on shifted frame difference to enhance the detection using spatial filters.

In [69] Herrero et al proposed a method of background subtraction using an extensive dataset of synthetically combined foreground & background objects. But this dataset does not cover challenges like scene illumination & distraction by shadows. In their experiments they account for indoor & outdoor environment. In comparison we account for various problems for background subtraction happening in the field of video surveillance system. Hence, we implemented motion detection & tracking system that not only enables comprehensive evaluation of background subtraction methods but also deals with the problem of illumination changes & dynamic background.

3 Database Collections

In our experiment we used two different databases. One is Weizmann database which is publically available and another is our own database this database contains 86 sequences having five & four of classes of actions (Jump, Run, Walk, Side, Handwaving, Jogging) respectively performed by 19 different subjects in two different conditions d1-d2.

d1 – Indoor Environment

d2 - Outdoor (Playground) + illumination variations

Few samples of Indoor environment

a) Run

b) Side Walk

c) Handwaving

d) Walk

e) Jump

Outdoor (Playground) + illumination variations Environment

a) Jogging

b) Handwaving

c) Walk

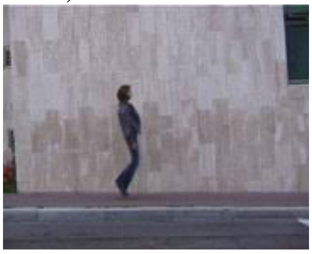
d) Jump

Fig.2: Sample of Own database

a) Run

b) Side Walk

c) Walk



d) Walk

Fig.3: Sample of Weizmann database

4 Methodology

Human motion perception may be divided into two Phases: first detection and, possibly, segmentation; then Tracking.

Motion detection & tracking algorithms have to cope with numerous challenges arising from the nature of video surveillance like gradual illumination changes, dynamic background, video noise and many others. To overcome these problems we proposed the system, consists of three frame differences and morphological blob analysis to avoid false motion detection. The flowchart of the proposed system is shown in Fig. 1.

The proposed work is divided into three phases. First phase consist of extracting three gray scale frames from the given video and then taking the frame difference between frame F_{n-2} and F_{n-1} and second frame difference between the frame F_{n-1} and current frame F_n. It is unlike other change detection algorithms where only the difference between the current and previous frame is considered. In second phase, morphological operation is performed on the resultant frame difference to suppress the remaining errors. Then the background regions are extracted, holes are filled and small regions are removed. We can get, two background masks, after morphological operations which are then compared with the threshold calculated followed by the AND operation applied to eliminate the false motion detection. Finally, the Kalman filter [14] is used to remove the noise and other changes in pixel due to illumination or any other reason. Extracted foreground is then tracked by the rectangle around it.

4.1 Gray Scaling

The input video taken from a video camera is colour image frame, it is first converted to gray image and each pixel is being represented by unsigned 8-bit integer. The conversion between RGB images to a gray scale image is given in (1)

$$G = 0.2989\,R + 0.5870G + 0.1140B \quad \text{..... (1)}$$

Where G is the gray scale image and R, G, B are the equivalent red, green and blue colour components of the image.

4.2 Dual Frame Differences

From the video three frames are extracted. So let, Frames are denoted by F and F_n, F_{n-1}, F_{n-2} represents the current, previous and second previous frames respectively. A simple background subtraction [13] yield the poor result and there are holes left. Therefore, to overcome the problem, three frame difference method is used [6]. In which two frame differences of three frames F_n, F_{n-1}, F_{n-2} are taken. One of F_{n-2}, F_{n-1} and second difference between F_n, F_{n-1}. These differences are labelled as Frame Subtraction 1 & Frame Subtraction 2.

4.3 Image Thresholding

These difference images show the "probability" of each pixel to be considered as a moving one. In this sense, we have to define a couple of thresholds T1 and T2. The result images will have 1's in pixels with motion, and 0's in pixels of static zones. This process is shown as:

$$\mathbf{IMT1}(x,y) = \begin{Bmatrix} \text{1 if Frame Sub1} \geq \text{T1} \\ \hline \text{0 if Otherwise} \end{Bmatrix} \quad \text{..... (2)}$$

$$\mathbf{IMT2}(x,y) = \begin{Bmatrix} \text{1 if Frame Sub2} \geq \text{C} \\ \hline \text{0 if Otherwise} \end{Bmatrix} \quad \text{..... (3)}$$

Where Frame Sub1 & Frame Sub2 are given by

$$FS_1 = F_{n-1} - F_{n-2} \quad ..... (4)$$

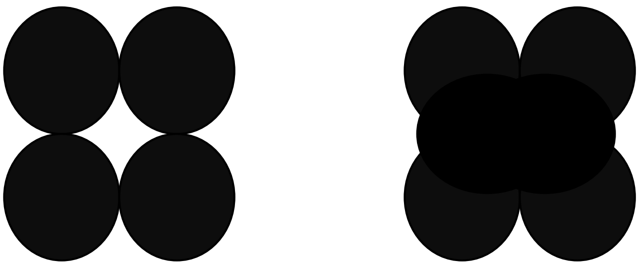
$$FS_2 = F_n - F_{n-1} \quad ..... (5)$$

These frames are having two coordinates x & y. Both the frames are of same size. We have selected $T1$ & $T2$ by Otsu's method [7].

4.4 Morphological Processing

There are many gaps and noise in output of dual frame subtraction method. Therefore, a morphological operation to decrease the noises and fill the gaps has been done. The morphological open operation is erosion followed by dilation, & morphological close operation is dilation followed by erosion using the same structuring element for both operations.

Background region is extracted, gaps are filled and small regions are removed. Opening generally smoothes the contour of an object, breaks narrow isthmuses, and eliminates thin protrusions. Closing generally fuses narrow breaks and long thin gulfs, and eliminates small holes and fills gap in the contour. Example of morphological processing is shown in Figure 4.



A) B)

Fig.4: A) Image before morphological processing.

B) Image after morphological processing.

4.5 AND Operation

AND operation is performed to find out the similarities between two difference images. After the morphological & AND operation the foreground pixels could be detected accurately. The equation for AND operation is given by

$$Hn = FS1 \wedge FS2 \quad ..... (6)$$

4.6 Kalman Filter

The output of the AND operation is still contaminated with different noise like white noise. As in our database both the camera and object were moving. The problem of correct detection of a target is difficult. Also, the database on which the algorithm was implemented was having the abrupt illumination changes. To overcome all these problems Kalman filter provides the solution.

The Kalman filter estimates the state of the dynamic system, even if the precise form of the system is unknown. The filter is very powerful in the sense that it supports the estimation of past, present and even the future state.

The Kalman filter [8] is recursive predictive filter that is based on the use of state space techniques and recursive algorithms. It estimates the state of dynamic system. In first step the state is predicted with the dynamic model and in the second step it is corrected with the observation model.

The dynamic model describes the transformation state over the time. It can be usually be represented by system of differential equation

$$\frac{X(\dot{t}) = d}{dt} X(t) = f(X(t), m(t)) \quad ..... (7)$$

For linear case this can be written as

$$X(\dot{t}) = FX(t) + n(t) \quad ..... (8)$$

Where F is the dynamic matrix and is constant, X(t) is the state vector, n(t) is the dynamic noise which is usually assumed to be white noise and has the covariance matrix Q(t).

The KF algorithm represents a realistic linear measurement update for the estimate and error covariance. KF can be structured into prediction and correction equations.

Prediction Equation is given by

$$\dot{X}(t) = \emptyset_0^t . \dot{X}(t_0) \quad ..... (9)$$

Where $\emptyset_0^t$ is the state transition matrix, which transforms any initial $X(t_0)$ to its corresponding state $X(t)$ at time *t*.

Correction Equation is given by

$$X^+(t_i) = X^-(t_i) + K(t_i).(l(t_i) - l^-(t_i)) \quad ..... (10)$$

In this equation the estimated state and the measurement state are weighted and combined to calculate the corrected state.

4.7 Target Tracking

Once filter operation is over, the detected foreground is fully visible from the background & there is less chance of misdetection of target. The fragmented target is represented through rectangular shape to envelope the target.

$$C_x = \frac{l}{2} \quad ..... (11)$$

$$C_y = \frac{b}{2} \quad \ldots\ldots (12)$$

Where l & b are the length & breadth of the detected foreground

$$l = X_{max} - X_{\min} \quad \ldots\ldots (12)$$

$$b = Y_{max} - Y_{\min} \quad \ldots\ldots (13)$$

$X_{\max}$, $X_{\min}$, Y_{max}, $Y_{\min}$ are the maximum and minimum spatial coordinates of the detected foreground region.

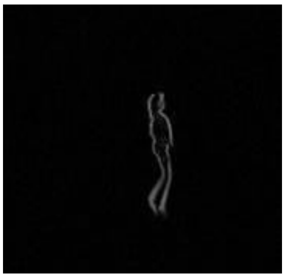
5 Experiment & Analysis

In this section, the algorithm has been implemented on the two different datasets i.e. Weizmann publically available & self made database. Weizmann database [180*144, Frame rate 25frames/sec] has four different actions Jump, Run, Walk and Side. All the actions are captured with steady camera. Our own database [1280*720, Frame rate 29frames/sec] is captured when both camera & target are moving in dynamic environment.

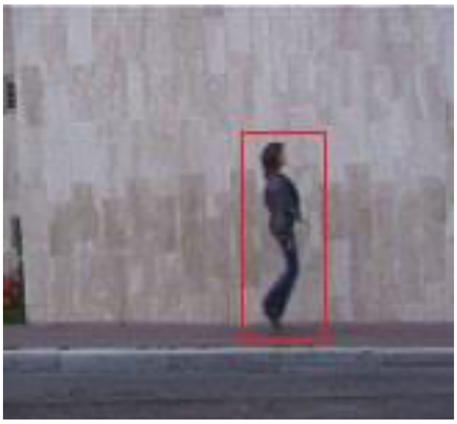
All the tests were run on Windows 7.0 platform. The PC is equipped with Pentium corei3 processor 2.13 GHz and 4 GB RAM Memory. Algorithms are implemented in MATLAB. The algorithm can detect and track moving object and is tested for steady & moving camera. Figure 5 shows the result for steady camera & Fig 6 & 7 shows result for moving camera & dynamic background.

From the result obtained fig. shown below it is clear that, algorithm achieves accurate target tracking [15] not only in steady but also in dynamic conditions. The test is performed for human detection & tracking.

a)



b)



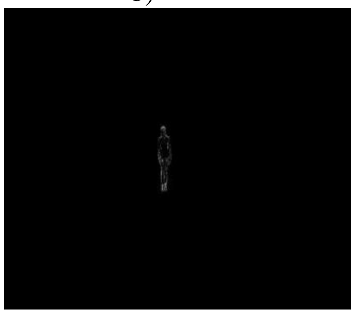
c)

Fig.5: a) 39th Video Frame of Weizmann Database (subject Jumping).
b) O/P of implemented Algorithm.
c) Tracked foreground.

a)

b)

c)



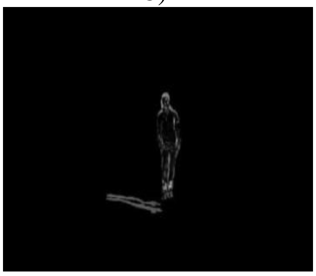
d)

e)

Fig.6: a) 7th Video Frame of Own Database.
b) Result of Background Subtraction method.
c) O/P of dual frame difference.
d) O/P of implemented Algorithm.
e) Tracked foreground.


a)

b)

c)

d)

e)

Fig. 7: a) 22ndVideo Frame of Own Database
b) Result of Background Subtraction method.
c) O/P of dual frame difference.
d) O/P of implemented Algorithm.
e) Tracked foreground.

The algorithm implemented has achieved highly robust & clean results in cluttered backgrounds and illumination change. This shows the effectiveness of algorithm.

6 Conclusion & Future Work

In this paper, we have implemented a precise method for detection and tracking of human motion to overcome the problem of unsteady camera and varying illumination condition. The experimental results carried out shows that the algorithm is efficient with all kind of video data set taken.

Our future extensions include further study on occlusion handling problems, which will be of interest to social studies and can be used for realistic crowd simulation. We will also focus on increasing the performance of the system by decreasing the time duration from frame to frame processing.

References:

1 D. Serby, E. K. Meier, and L. V. Gool, "Probabilistic Object Tracking Using Multiple Features, "IEEE Proc. Of International Conf on Pattern Recognition Intelligent Transportation Systems, Vol. 6, pp.43-53, March 2004.

2 L. Li, S. Ranganath, H. Weimin, and K. Sengupta, "Framework for Realtime Behavior Interpretation Form Traffic Video ", IEEE Tran. On Intelligen Transportation Systems, March 2005, Vol.6, No.1, pp.43-53.

3 P. Kumar, H. Weimin, I.U. Gu, and Q. Tian, "Statistical Modeling of Complex Backgrounds for Foreground Object Detection", IEEE Trans.on Image Processing,Vol.13,Mo.11, pp.43-53,November 2004.

4 Z. Zivkovi, "Improving the selection of feature points for tracking", In pattern Analysis and Applicarions, vol.7, no .2, Copyright Springer Verlag London Limited, 2004.

5 J. Lou, T.Tan, W. Hu, H. Yang,and S. H. Maybank, "3D Model-based Vehicle Tracking", IEEE Trans.on Image Processing, Vol. 14, pp. 1561-1569, October2005.

6 Jun Yang, Tusheng Lin, Bi Li, “Dual Frame Differences Based Background Extraction Algorithm”, ICCP2011 Proceedings, pp. 44-47, 2011.

7 N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, pp. 62-66, 1979.

8 Rachel kleinbauer, “Kalman Filtering Implementation with MatLab”, University Stuttgart Institute of Geodesy, Helsinki Nov, 2004.

9 Jun Yang, Tusheng Lin, Bi Li, Member, IEEE, “Dual Frame Differences Based Background Extraction Algorithm”, ICCP Proceedings, 2011.

10 D. Serby, E. K. Meier, and L. V. Gool, "Probabilistic Object Tracking Using Multiple Features, "IEEE Proc. Of International Conf on Pattern Recognition Intelligent Transportation Systems, vol. 6, pp. 43-53, March 2004.

11 E. Durucan and T. Ebrahimi. Moving object detection between multiple and color images. In

IEEE Conf. Advanced Video and Signal Based Surveillance, pages 243–251, 2003.

12 Zalili Musa & Junzo Watada; “Multi-camera Tracking System in Large Area Case,” in IEEE at Korea 2009.

13 Y. Dhome, N. Tronson, A. Vacavant, T. Chateau, C. Gabard, Y. Goyat, and D. Gruyer. A Benchmark for Background Subtraction Algorithms in Monocular Vision: a Comparative Study. In International Conference on Image Processing Theory, Tools and Applications (IPTA 2010). 7–10 July 2010.

14 Li Ying-hong, Pang Yi-gui, Li Zheng-xi and Liu Ya-li, “An intelligent tracking technology based on kalman and mean shift algorithm,” IEEE Second Intl. Conf. on Computer Modelling and Simulation 2010.

15 Z. Zivkovi, "Improving the selection of feature points for tracking”, In pattern Analysis and Applicarions, vol.7, no .2, Copyright Springer Verlag London Limited, 2004.

16 P. Perez, C. Hue, J. Vermaak, and M. Gangnet. "Color-Based Probabilistic Tracking", in *Proceedings of the European Conference on Computer Vision* (ECCV), 2002.

17 F. Porikli, O. Tuzel, and P. Meer. ``Covariance tracking using model update based on Lie algebra", in *in Proceedings of the Computer Vision and Pattern Recognition* (CVPR), 2006.

18 Xiaoqin Zhang, Weiming Hu, Zixiang Zhao, Yan-guo Wang, Xi Li and Qingdi Wei,”SVD based kalman particle filter for robust visual tracking,” IEEE 2008, 978-1-4244-2175-6.

19 R. T. Collins, A. J. Lipton, and T. Kanade, “Introduction to the special section on video surveillance,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, pp. 745–746, Aug. 2000.

20 J. Steffens, E. Elagin, and H. Neven, “Person spotter-fast and robust system for human detection, tracking and recognition,” in *Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition*, 1998, pp. 516–521.

21 R. T. Collins, A. J. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt, and L.Wixson, “A system for video surveillance and monitoring,” Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep., CMU-RI-TR-00-12, 2000.

22 I. Haritaoglu, D. Harwood, and L. S. Davis, “W: Real-time surveillance of people and their activities,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, pp. 809–830, Aug. 2000.

23 J. Fernyhough, A. G. Cohn, and D. C. Hogg, “Constructing qualitative event models automatically from video input,” *Image Vis. Comput.*, vol. 18, no. 9, pp. 81–103, 2000.

24 Roland Kehl, Luc Van Gool, Markerless tracking of complex human motions from multiple views, Computer Vision and Image Understanding (CVIU) 104 (2–3) (2006) 190–209.

25 Thomas B. Moeslund, Erik Granum, A survey of computer vision based human motion capture, Computer Vision and Image Understanding (CVIU) 81 (3) (2001) 231–268.

26 Vladimir I. Pavlovic´, James M. Rehg, Tat-Jen Cham, Kevin P.Murphy, A dynamic Bayesian network approach to figure tracking using learned dynamic models, in: Proceedings of the International Conference on Computer Vision (ICCV’99), vol. 1, Kerkyra, Greece, September 1999, pp. 94–101.

27 J.Y. Zou, Q. Pan, Y. Liang, H.C. Zhang, Y.M. Cheng, “Model Switching Based Adaptive Background Mdeling Approach”, *Journal of Acta Automatica Sinica,*, 2007 May, 33(5): pp. 467-473.

28 N. Friedman, S. Russell, “Image segmentation in video sequences: a probabilistic approach”, *In: Proceeding of Thirteenth Conference on Uncertainty in Artificial Intelligence*, Providence, Rhode Island, USA, Morgan Kaufmann Publishers, 1997, pp. 175-181.

29 C.R. Wren, A. Azarbayejani, T. Darrell, A. Penyland, “Pfinder: Real-time tracking of the human body”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): pp. 780-785.

30 Stauffer C, Grimson W. Adaptive background mixture models for real-time tracking. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Fort Collins, Colorado, USA, IEEE, 1999: 245-252.

31 Z. Zivkovic, “Improved adaptive Gaussian mixture model for background subtraction”, *In: Proceedings of the 17th International Conference on Patter Recognition*, Cambridge, United Kingdom, IEEE, 2004, 2: pp. 28-31.

32 D. A. Ross, J. Lim, R. Lin, and M. Yang. Incremental learning for robust visual tracking. *IJCV*, 77(1-3):125–141, 2008.

33 B. Babenko, M. Yang, and S. Belongie. Visual tracking with online multiple instance learning. *CVPR*, 2009.

34 J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, “Robust face recognition via sparse representation”. PAMI, 31(2):210–227, 2009.

35 X. Mei and H. Ling, "Robust visual tracking using l1 minimization". In ICCV, 2009.
36 B. Liu, L. Yang, J. Huang, P. Meer, L. Gong, and C. Kulikowski, "Robust and fast collaborative tracking with two stage sparse optimization". In ECCV, pages 624–637, 2010.
37 R. Poppe, "Vision-based human motion analysis: An over-view," *Comput. Vis. Image Understanding*, vol. 108, pp. 4–18, Oct. 2007.
38 C. Cedras and M. Shah, "A survey of motion analysis from moving light displays," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 1994, pp. 214–221.
39 C. C´edras and M. Shah, "Motion-based recognition: A survey," *Image Vision Comput.*, vol. 13, pp. 129–155, 1995.
40 A. F. Bobick, "Movement, activity and action: The role of knowledge in the perception of motion," *Philos. Trans. Roy. Soc. B, Biol. Sci.*, vol. 352, pp. 1257–1265, 1997.
41 S. Avidan. Support vector tracking. In CVPR, pages 184– 191, 2001.
42 R. T. Collins, Y. Liu, and M. Leordeanu. Online selection of discriminative tracking features. PAMI, 27(10):1631–1643, 2005.
43 X. Xie and K. Lam. An Efficient Illumination Compensation Scheme for Face Recognition. Proc. Control, Automation, Robotics and Vision Conference: Vol. 2. 1240 – 1243, 2004.
44 J. Cai, M. Shehata, W. Badawy, and M. Pervez. An Algorithm to Compensate for Road Illumination Changes for AID Systems. Proc. IEEE Intelligent Transportation Systems Conference: 980 – 985, 2007.
45 A. D. Jepson, D. J. Fleet, and T. F. El-Maraghi. "Robust online appearance models for visual tracking", *PAMI*, 25:1296-1311, 2003.
46 D. Freedman and M. W. Turek. "Illumination-Invariant Tracking via Graph Cuts", *CVPR*, 10-17, 2005.
47 C. Yang, R. Duraiswami, and L. S. Davis. "Fast multiple object tracking via a hierarchical particle filter", *ICCV*, 212-219, 2005.
48 A. Yilmaz, O. Javed, and M. Shah. 'Object tracking: A survey". *ACM Comput. Surv.* 38(4), 2006.
49 Q. Yu, T. B. Dinh and G. Medioni. "Online tracking and reacquisition using co-trained generative and discriminative trackers", *ECCV*, 678- 691, 2008.
50 Christian Wojek, Kai Nickel, Rainer Stiefelhagen; "Activity Recognition and Room Level Tracking" in the IEEE International Conference on Multi sensor Fusion and Integration for Intelligent Systems, 2006, Heidelberg, Germany;pp 1-6.
51 E.h.joraba, C.o.urunuela, j.senar, "Detected motion classification with a double-background and a neighborhood-based difference", Elsevier science, Pattern Recognition Letters, Volume 24, pp. 2079 - 2092 , August 2003.
52 G. Welch, G. Bishop, "An Introduction to the Kalman Filter", ACM Inc., 2001.
53 C.R. Wren, A. Azarbayejani, T. Darrell and A.P. Pentland, "Pfinder: real-time tracking of the human body," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, pp. 780-785, 1997.
54 M.J. Hossain, J. Lee, and O. Chae. An Adaptive Video Surveillance Approach for Dynamic Environment. Proc. Int'l Symposium on Intelligent Signal Processing and
55 Y. Wang; T. Tan, and K. Loe. A probabilistic method for foreground and shadow segmentation. Proc. Int'l Conference on Image Processing: Vol. 3. 937 – 940, 2003.
56 B.K.P. Horn. Robot Vision. (McGraw-Hill, New York, NY, U.S., 1986), 185 – 216, 1986.
57 G.D. Hager and P.N. Belhumeur. Real-time tracking of image regions with changes in geometry and illumination. Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition: 403 – 410, 1996.
58 H. Wu, P. Jiang and J. Zhu. An Illumination Compensation Method for Images under Variable Lighting Condition. Proc. IEEE Conference on Robotics, Automation and Mechatronics: 1022 – 1026, 2008.
59 Kristen Grauman, Gregory Shakhnarovich, Trevor Darrell, Inferring 3D structure with a statistical image-based shape model, in: Proceedings of the International Conference on Computer Vision (ICCV'03), vol. 1, Nice, France, October 2003, pp. 641–647.
60 W. H. Press, Saul A. Teukolsky, William T. Vetterling and Brian P. Flannery, "Numerical Recipes in C: The art of Scientific Computing", Second Edition, Cambridge University press, 2002.
61 L. Wang, W. Hu, and T. Tan, "Recent developments in human motion analysis," Pattern Recognition, vol. 36, pp. 585-601, 2003.
62 P. Kumar, A. Mittal, and P. Kumar, "Study of Robust and Intelligent Surveillance in Visible and Multi-modal Framework," Informatica 32, pp. 63-77, 2008.
63 Y. Ivanov and A. Bobick, "Recognition of Visual Activities and Interactions by Stochastic

Parsing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 8, pp. 852-872, 2000.

64 Isaac, C. Medioni, "Detecting and tracking moving objects for video surveillance", in Proceedings of ICPR Fort Collins CO. Jun. 23-25, 1999.

65 H. Li, S. Wu, S. Ba, S. Lin and Y. Zhang, "Automatic detection and recognition of athlete actions in diving video," *Proc. Int. Conf. Multimedia Modeling*, LNCS 4352, pp. 73-82, 2007.

66 R. Fablet and M.J. Black, "Automatic detection and tracking of human motion with a view-based representation," in *ECCV*, 2002, vol. 1, pp. 476–491.

67 H. Sidenbladh, "Detecting human motion with support vector machines," in *ICPR*, 2004, vol. 2, pp. 188–191.

68 P. Viola, M. Jones and D. Snow, "Detecting pedestrians using patterns of motion and appearance," *ICCV*, vol. 02, pp. 734–741, 2003.

69 S. Herrero and J. Besc´os. Background subtraction techniques: Systematic evaluation and comparative analysis. In Advanced Concepts for Intelligent Vision Systems, pages 33–42.Springer, 2009.

70 D. Parks and S. Fels. Evaluation of background subtraction algorithms with post-processing. In Proc. of IEEE Int. Conf. on Advanced Video and Signal Based Surveillance, pages 192–199, 2008.

71 Li, L.; Huang, W.; Gu, I.; Tian, Q. Statistical modeling of complex backgrounds for foreground object detection. IEEE Trans. Image Process. 2004, 13, 1459–1472.

72 Ismail Haritaoglu, David Harwood and Larry S. Davis, "W4: Real-Time Surveillance of people and their Activities", *Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 22, No. 8, pp 809-830, August 2000.