INTERNATIONAL JOURNAL OF RESEARCHES IN BIOSCIENCES, AGRICULTURE AND TECHNOLOGY

www.vmsindia.org

PERFORMANCE ANALYSIS OF ADABOOST CLASSIFIER AND OPTICAL FLOW METHOD ALGORITHMS FOR HUMAN MOTION DETECTION

Moiz A. Hussain[1] and G. U. Kharat[2]

[1]ENTC Department, Anuradha Engineering College, Chikhli, India

[2]Sharadchandra Pawar College of Engineering, Otur, India

mymoiz2004@yahoo.co.in

Abstract

An algorithmic approach to detect human motion for video surveillance system is discussed in this paper. The considerable study in this area has been encouraged by the fact that many application areas, including surveillance, Human–Computer Interaction and automatic annotation, will gain from a robust solution. In this paper, a framework for the human motion tracking using two algorithms is presented, first, using AdaBoost classifier. The algorithm consists of three steps: at first, a multiscale image features extracted from an image is used. At the second stage, extracted features are represented by the sparse matrix representation. These sparse matrix represented values are classified using AdaBoost classifier for motion tracking applications. Second using optical flow method where optical flow values are extracted which is then processed by illumination insensitive tracker. Experimental result shows that both the method gives better performance as compared to the other state of the arts methods of video surveillance.

Keywords—Video surveliannce; Sparse Matrix; AdaBoost Classifier.

Introduction

A fundamental problem for all automatic video surveillance systems is to detect objects of interest in a given scene. A usually used technique for segmentation of moving objects is background subtraction **[1]**. In-spite, of tremendous research in the field of surveillance system still it is a difficult task to build up efficient models for object tracking due to factors such as illumination change, pose variation, occlusion, and motion blur. Many tracking algorithms frequently revise models with samples from observations in recent frames. Despite much success has been demonstrated, numerous issues remain to be addressed. First, many tracking algorithms often encounter the drift problems. Second, while these models are data dependent, there does not exist sufficient amount of data to learn at the outset. As an effect of self-taught training, misaligned samples are likely to be added and degrade the appearance models. A classic surveillance application has 3 constructional blocks: motion detection, object tracking and advanced 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 grouped into 3 major classes: shape based models **[2]**, area based models **[3, 4]**, and characteristic point-based models **[5, 6]**. Also most of the programmed surveillance system uses optical flow method for object detection. This works on the principle that Optical flow **[7, 8]** cannot be calculated locally, because merely one independent dimension is accessible from the image sequence at a point, whereas the flow rate has two factors a second constraint is essential. Smart programmed tracking system should be able to sense modification caused by a new object, whereas moving background regions, such as clouds, rain or a flag waving in the wind, should be recognized as an element of the background. Optical flow; It computes an independent estimate of motion at each pixel; this usually involves reducing the intensity among consequent pixels summed over the image. It is supposed that these variations are because of motion and not due to other effect, for example lightening effect. The illumination can vary in an interior and an outside image sequence. The illumination in an interior sequence can vary as the radiance varies as a minimum 3,600 times per minute by reason of indoor lighting; on the other hand, the image taken by a camera is 25 ~ 30 per second. As a result, illumination has a consequence on the captured image sequence. The elucidation in an open-air sequence also changes for a variety of reasons for example the sun light change & this variations are extremely quick compared to indoor sequence.

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 the system, compared to the other state-of-the-art methods is that it reduces the number of false detections, as multiscale feature matrix is formed using sparse representation & pixels are the correctly classified using AdaBoost classifier. The system proposed is a solution to the problem of efficiently detecting & tracking moving object with easy form and less processing time. One of the merits of this method is that it decreases the number of wrong detections, as pixel-level classification can be carried out in regions with considerable motion only.

The system flow diagram of proposed model is shown in Figure 1. Figure shows the different stages involved in the proposed system. First, image frames are extracted from video sequence, and then, multiscale image features are calculated by using sparse matrix. Extracted feature vectors are then given to the AdaBoost classifier for the tracking purpose.

In the second approach, to detect foreground objects, first, optical flow algorithm is applied. This is then combined with the illumination insensitive template matching method to accurately track the object for visual surveillance system. The system proposed is a solution to the problem of efficiently detection and tracking of moving people with ease and less processing time. Figure 2 shows the different stages involved in the proposed system. First, image frames are taken out from video, and then, optical flow calculation is done as described in section 3.2.7. Extracted feature of optical flow calculation; is then combine with the illumination insensitivity hyperplane method **[172]** to perfectly track the object.

The remainder of this paper is organized as follows: Section 2 discusses some of the related methods. Section 3, presents the system methodology in detail. Section 4, describes the database collection for the experimentation work. Section 5 presents some experimental results & Finally, Section 6 concludes the paper.

RELATED WORK

There is an enormous literature on video surveillance for human motion detection & tracking. In our recent work **[9]**, an adaptive threshold initialization system to segment objects from a video based on the supposition that the moving objects are visually separated without overlapping regions is presented. Avidan **[16]** uses a neural network i.e. Support Vector Machine (SVM) classifier offline and continue it within the optical flow framework for object tracking. Collins et al. **[21]** use inconsistency ratio of interested object & background classes to decide discriminative characteristics for object tracking.

Bao et al. **[22]**, and Ravi et al. **[28]**, have proposed methods to forecast few actions like walking, standing, running, sit-ups, and others using features from raw accelerometer data and a variety of different learning algorithms. On the other hand, they do not use this information for indoor positioning. Because this method concentrates on intra-image lighting variations & spatial compared to sequential, inter image lighting changes, it is a dissimilar group of lighting compensation. The research by Cwojek et. al **[29]** presented a method for several human action identification in an office situation with immediate tracking. In this work, video and audio features are implemented to use a multilevel Hidden Markov Model (HMM) framework. A Pfinder **[34]** technique was developed to explain a moving person in an indoor environment. It tracks a solo non-overlapped human being in complex scenes. A combination model of 3 factors with an online EM algorithm is developed to model the appearance changes during tracking **[35]**. The graph-cut algorithm was also developed for detection & tracking by calculating a lighting invariant optical flow field **[36]**. Ross et al. **[41]** 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 **[42]** successfully tracks an object in real time where lighting conditions change and the object is occluded by others.

The recent development of sparse representation **[43]** has attracted considerable interest in object tracking **[44, 45]** 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. We implemented motion detection & tracking system that not only enables comprehensive evaluation of sparse matrix representation methods but also deals with the problem of illumination changes & dynamic background.

Methodology

Method 1: A simple yet effective and efficient tracking algorithm based on features extracted from a multiscale image feature space with data-independent basis is proposed. The proposed model employs non-adaptive random projections that preserve the structure of the image feature space of objects. A sparse measurement matrix is constructed to efficiently extract the features for the appearance model. Sample images are compressed of the foreground target and the background using the same sparse measurement matrix. The tracking task is formulated as a binary classification via a AdaBoost classifier with update in the compressed domain. A coarse-to-fine search strategy is adopted to further reduce the computational complexity in the detection procedure.

A. Image Representation

From the video the frames are extracted. The multiscale image representation is formed by convolving the input image with a Gaussian filter of different spatial variances **[173]**. For each sample image its multiscale representation is constructed by convolving image with a set of filters at multiscale $\{F_{1,1} \ldots . F_{w,h}\}$ defined by

$$F_{w,h}(x, y) = \frac{1}{wh} \times \begin{cases} 1, & 1 \le x \le w, 1 \le y \le h \\ 0, & Otherwise \end{cases}$$

Where, w and h are the width and height of rectangle filter, respectively. Each filtered image is represented as a column vector in R^{wh} and concatenate these vectors as a very high-dimensional multiscale image feature vector $x = (x\_1, \ldots . x\_m)^T \in R^m$ where $m = (wh)^2$. The dimensionality m is typically in the order of 10^6 to 10^{10}.

B. Integral Image

Rectangle features can be computed very rapidly using an intermediate representation for the image which is called the integral image. The integral image at location x,y contains the sum of the pixels above and to the left of x,y inclusive:

$$ii(x, y) = \sum_{x' \le x, y' \le y} i(x^i, y')$$

Where, $ii(x, y)$ is the integral image and $i(x, y)$ is the original image. Using the following pair of recurrences:

$$s(x, y) = s(x, y - 1) + i(x, y)$$

$$ii(x, y) = ii(x - 1, y) + s(x, y)$$

Where, $s(x, y)$ is the cumulative row sum, $s(x, -1) = 0$, and $ii(-1, y) = 0$ the integral image can be computed in one pass over the original image.

C. Image Features

The main purpose of using features instead of raw pixel values as the input to a learning algorithm is to reduce the in-class while increasing the out-of-class variability compared to the raw data and thus making classification easier. Features usually encode knowledge about the domain, which is difficult to learn from the raw and finite set of input data. A very large and general pool of simple haar-like features combined with feature selection therefore can increase the capacity of the learning algorithm. The speed of feature evaluation is also a very important aspect since almost all object detection algorithms slide a fixed-size window at all scales over the input image. then the set of all possible features of the form

$$feature_1 = \sum_{i \in I = \{1, \ldots N\}} w_i \cdot RecSum(r_i)$$

Where the weights $w_{i \in} R$, the rectangles r_i and N are arbitrarily chose.

D. Fast Feature Computation

All the features can be computed very fast and in constant time for any size by means of two auxiliary images. For upright rectangles the auxiliary image is the Summed Area Table $SAT = (x, y) \cdot SAT(x, y)$ is defined as the sum of the pixels of the upright rectangle ranging from the top left corner at (0, 0) to the bottom right corner at (x, y).

$$SAT(x, y) = \sum_{x' \le x, y' \le y} I(x', y')$$

E. Classifier Construction

It is assumed that all the elements in v are independently distributed and model them with an AdaBoost learning algorithm **[49]**. Given a feature set and a training set of images, any number of machine learning approaches could be used to learn a classification function. In our system a variant of AdaBoost is used both to select a small set of features and train the classifier. In its original form, the AdaBoost learning algorithm is used to boost the classification performance of a simple (sometimes called weak) learning algorithm. There are a number of formal guarantees provided by the AdaBoost learning procedure.

For each feature, the weak learner determines the optimal threshold classification function, such that the minimum number of examples is misclassified. A weak classifier $Cj(x)$ thus consists of a feature fj, a threshold θj and a polarity pj indicating the direction of the inequality sign:

$$Cj(x) = \begin{cases} 1 & \text{if } pjfj(x) < pj\theta j \\ 0 & \text{otherwise} \end{cases}$$

In terms of computation, this classifier is probably faster than any other system.

Unfortunately, the most straight forward technique for improving detection performance, adding features to the classifier, directly increases computation time.

Method 2: Human motion detection & tracking algorithms have to deal with numerous problems occurring from the nature of video surveillance like gradual illumination changes, dynamic background, video noise and many others. Our objective is to provide an efficient method to find the moving objects in the video frame as well as to track them. The algorithm to learn a system for motion detection & tracking can be briefly quoted as follows.

1. Extract the frames from the given video sequence.
2. Take last & current frame and calculate the optical flow in between the 2 frames.
3. Compute the features as optical flow output.
4. Estimate µ from the template of an image.
5. Compute the value of (βt, ɣt) for illumination insensitive iteration cycle.
6. Combine the value of step 5 and step 3 to track the given object in a video sequence.

In order to track an object in a video sequence optical flow constraint equation can be used.

$$Ix\,(x_i,y_i)\cdot dx\,(x_i,y_i) + Iy\,(x_i,y_i)\cdot dx\,(x_i,y_i) = -I_t\,(x_i,y_i)$$

IV. DATABASE COLLECTIONS

In our experiment two different databases is used. 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

TABLE 1: EXPERIMENT CONDITIONS

Sequence	1280 x720 pixels, 24 bit color 30 fps,
CPU	Corei3, 2.13GHz
RAM	4GB
OS	Windows 7.0
MATLAB	Version 7.10.0.499 (R2010a)

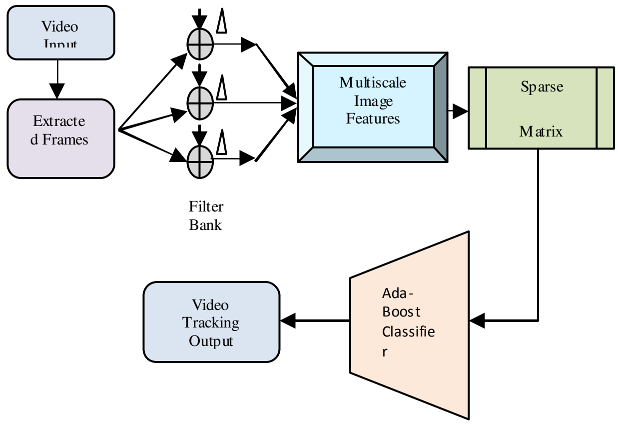


Figure 1: Adaboost System Overview

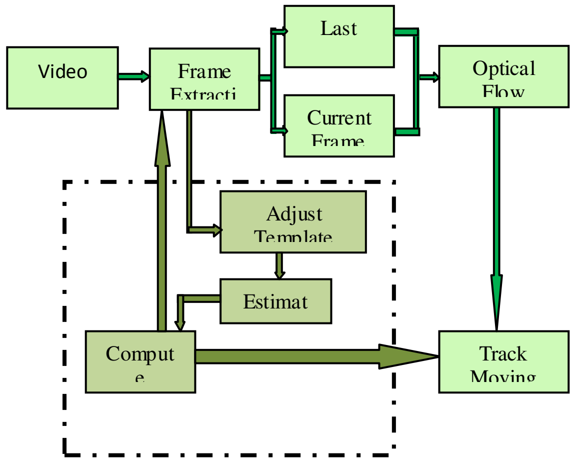
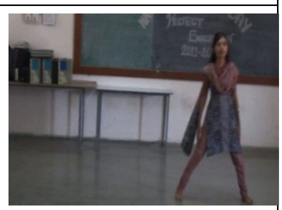
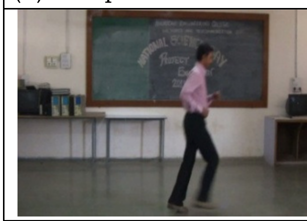


Figure 2: Optical System Overview

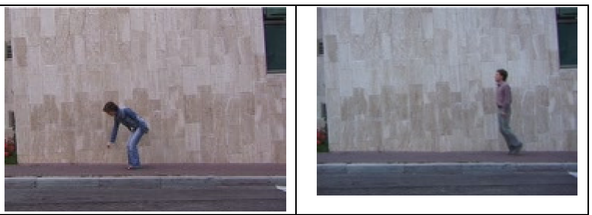
Few samples of Indoor Environment	
	
(A) Jump	(B) Sidewalk
	
(C) Run	(D)Walk
Outdoor (Playground) + Illumination Variations Environment	
	

(A) Jogging	(B) Jump

(C)Walk	(D) Handwaving

(C) Run	(D) Skip

Figure 3: Sample of Own & Weizmann database



(A) Bend	(B) Jump

Frame # 1	Frame # 18	Frame # 29	Frame # 41	Frame #5 6	Frame # 67	Frame # 69

Basketball sequence from CMU LAB Motion database

Frame # 1	Frame # 13	Frame # 20	Frame # 26	Frame # 35	Frame # 39	Frame # 44

AAK_Walk_d2 sequence from Own database (outdoor sequence)

Frame # 1	Frame # 11	Frame # 17	Frame # 27	Frame # 33	Frame # 40	Frame # 44

Ira_bend sequence from Weizmann database

Frame # 1	Frame # 4	Frame # 9	Frame # 14	Frame # 20	Frame # 24	Frame # 30

PAU_Walk_d1 sequence from Own database

Figure 5: Snapshots of experimental results (Indoor sequence)

Result

The proposed method is compared with the general sparse representation method for the condition of detection precision and computational time. The experimental conditions are summarized in Table I. The sparse matrix representation method uses feature selection of the object. To improve final estimation of the moving object AdaBoost classifier is trained with the extracted multiscale features and the moving object is accurately tracked. The proposed method gives the best result figure 5 compared to the other state of the art method in terms of preciseness and computational time.

Conclusion

In this paper, a robust & efficient multiscale sparse matrix features based object tracking algorithm that uses the reliable sparse nature of candidate particle representations using a glossary of object and background templates is proposed. A visual tracking system is modelled as a classification problem that is regularized by AdaBoost classifier at the level of particles, and presents an efficient solution. The performance of tracking algorithms against a number of competing state-of-the-art methods on different challenging image sequences is analyzed. Qualitative and quantitative experimental results show that the proposed tracking algorithms outperform state-of-the- art methods, especially in the presence of pose variations, illumination changes, and abrupt motion.

References

Lee B., Hedley M., "Background Estimation for video Surveillance," IVCNZ 2002, vol. 1, pp. 315-350,2002.

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.

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.

N. A. OGALE, *A survey of techniques for human detection from video,* http://www.cs.umd.edu/scholarlypapers/neeti Paper.pdf, Department of Computer Science, University of Maryland, College Park.

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.

[8] 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.

I. Austvoll. A Study of the Yosemite Sequence Used as a Test Sequence for Estimation of Optical Flow, pages 659– 668. Lecture Notes in Computer Science. Springer Berlin/ Heidelberg, 2005.

L. Alvarez, R. Deriche, T. Papadopoulo, and J. S´anchez. Symmetrical dense optical flow estimation with occlusions detection. In *Proc. ECCV*, pages 721–735, 2002.

Moiz A. Hussain, Dr. G.U.kharat, "Video Based Human Detection System Using Movement Analysis" IEEE, International conference on convergence of technology, I2CT-2014 .

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.

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.

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.

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.

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.

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.

S. Avidan. Support vector tracking. In CVPR, pages 184– 191, 2001.

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

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

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

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.

R.T. Collins, A.J. Lipton, T. Kanade, Introduction to the special section on video surveillance, IEEE Trans. on Pattern Analysis and Machine Intelligence, 22 (8) (2000) 745-746.

S. S. Bao and L. Intille, *Activity Recognition from User-Annotated Acceleration Data.* New York: Springer-Verlag, 2004, pp. 1–17.X.

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.

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

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

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.

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.

N. Ravi,N.Dandekar, P. Mysore, and M. L. Littman, “Activity recognition from accelerometer data,” in *Proc. 17th Conf. Innovat. Appl. Artif. Intell.*, 2005, pp. 1541–1546.

Christian Wojek, Kai Nickel, Rainer Stiefelhagen; “Activity Recognition and Room Level Tracking” IEEE International Conference on Multi sensor Fusion and Integration for Intelligent Systems, Heidelberg, Germany; pp 1-6, 2006.

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

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.

B.K.P. Horn. Robot Vision. (McGraw-Hill, New York, NY, U.S., 1986), 185 – 216, 1986.

Jin, H., Favaro, P., Soatto, S.: Real-Time Feature Tracking and Outlier Rejection with Changes in Illumination. In: International Conference on Computer Vision, Los Alamitos, IEEE Computer Society (2001) 684–689.

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.

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.

A. D. Jepson, D. J. Fleet, and T. F. El-Maraghi. “Robust online appearance models for visual tracking”, *PAMI*, 25:1296-1311, 2003.

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.

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

Y. Dhome, N. Tronson, “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.

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.

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

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

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.

X. Mei and H. Ling, “Robust visual tracking using l1 minimization”. In ICCV, 2009.

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.

R. Poppe, “Vision-based human motion analysis: An overview,” *Comput. Vis. Image Understanding*, vol. 108, pp. 4–18, Oct. 2007.

R. Poppe, “Vision-based human motion analysis: An overview,” *Comput. Vis. Image Understanding*, vol. 108, pp. 4–18, Oct. 2007.

Rainer Lienhart and Jochen Maydt, “An Extended Set of Haar-like Features for Rapid Object Detection”, Intel Labs, Intel Corporation, Santa Clara, CA 95052, USA.