Implementation of Real-time segmentation mechanism of moving regions in image sequences

Rekha Tiwari (M-Tech Scholar)
rekhatiwar@gmail.com
Sumit Dubey
Assistant Professor ECE Department JNCT, Rewa, India
Sneha Singh
Singhsneha1888@gmail.com
HOD, ECE, JNCT, Rewa, India

Abstract
The state of the art of all the motion detection techniques that are being employed currently in real time motion sensing systems. The work discusses mechanisms like Optical Flow, Frame differencing, Running Gaussian average as viable mechanisms for motion sensing and also their limitations. The work primarily focuses on Adaptive Gaussian Mixture Models as the preferred mechanism over other options. The work presents simulated results of motion sensing under different ambient conditions and different illumination conditions. The work presents a comparative analysis of all the existing mechanisms with the proposed mechanism.

Keywords: Motion Sensing, Background Subtraction, AGMM, GMM

Introduction
It is the human desire that has led to automatic detection systems and intelligent surveillance systems which make lives easier as well as enable us to compete with tomorrow's technology. On the other hand it has pushed us to analyze the challenge sin the field of automated video surveillance in light of the advanced artificial intelligence systems.

The surveillance cameras nowadays are already prevalent in secured commercial locations, with camera outputs being recorded to tapes that are either rewritten or periodically stored in video archiving systems. In order to benefit from this prerecorded digital data, detecting any moving object from the scene is required and that too without engaging any human aid. Real-time segmentation of moving regions in image sequences has been a fundamental step in many vision systems.

Motion Detection
Motion detection in consequent images the detection of the true moving object in the scene. In real time video surveillance systems, motion detection refers to the capability of the system to detect motion and capture the events and time of occurrence. That also

requires a software-based monitoring algorithm which in turn will signal the surveillance camera to begin
capturing the event when motion activity is detected. This is also called activity detection. An advanced motion detection surveillance system can analyze the type of motion for triggering an alarm system. In this project, however, the work confines to the robust sensing of activity in prerecorded video feed possibly taken from an associated real time surveillance mechanism and its associated mechanisms (morphological operations, filtering, shadow removal etc.) which in turn can be associated with a hardware based surveillance system. However, the development of that is not the scope of this work.

Problems and Issues
1. Optical Flow and Image Motion
2. Occluding Surfaces and Independently Moving Objects
3. Transparency
4. Prefiltering and Differentiation

PROBLEM STATEMENT
Earlier work of motion sensing and detection we faced many problem and also all the background models discussed so far have many limitations;
1. They ignore any correlation between neighbouring pixels.
2. The rate of adaption may not match the moving speed of the foreground objects.
3. Non-Stationary pixels from moving leaves or shadow cast by moving objects are easily mistaken as true foreground objects.
4. System will not cope with moderate change in light levels.
5. Lack of size discrimination means compromise in setting up.

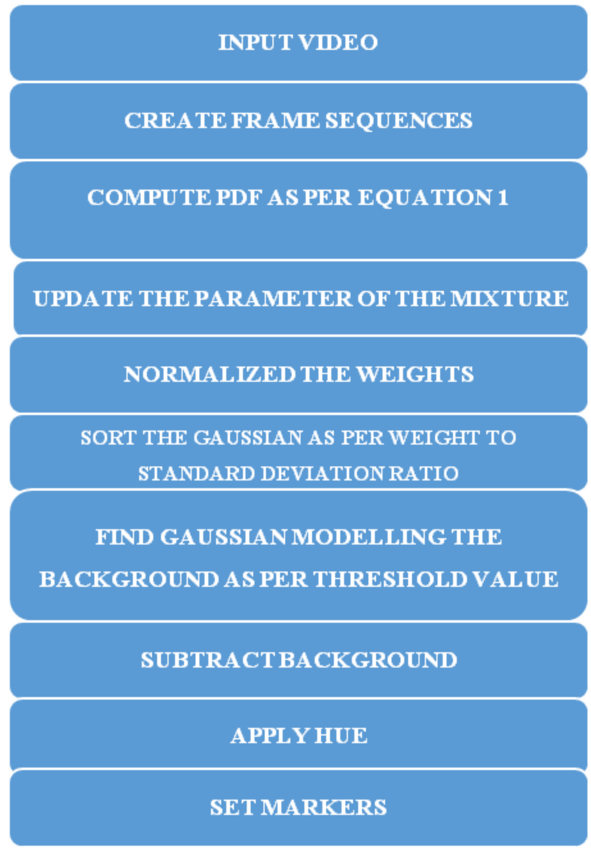
6. Non-Uniform sensitivity with range.
7. System will not cope with size variation due to perspective.
8. Slow processing speed can miss active object.
9. Inability to discriminate between small high contrast dark and large low contrast objects.
10. System cannot distinguish between a person moving in a line and waving object.
11. Single processor increase time between frame comparisons.
12. If an object close to the camera would activate far more cells than a person in the background. Simple cell count system may offer some improvement in false detection but do not offer accurate size discrimination.

These limitations or problems are eliminated in our proposed work by using Adaptive Gaussian Mixture Model.

PROPOSED METHODLOGY Background modeling by Gaussian mixtures is a pixel based process. Let x be a random process representing the value of a given pixel in time. A convenient framework to model the probability density function of x is the parametric Gaussian mixture model where the density is composed of a sum of Gaussians. Let p(x) denotes the probability density function of a Gaussian mixture comprising K component densities.

$$p(X) = \sum_{k=1}^{K} \omega_k N(X; \mu_k, \sigma_k) \qquad \text{Eqn.1}$$

Where ω_k are the weights and N(x; μ_k, σ_k) is the normal density of mean $\mu_k$and covariance matrix $\Sigma_k = \sigma_k I$, (I denotes the identity matrix). The mixture of Gaussians algorithm, proposed by Stauffer and Grimson [12] estimates these parameters over time to obtain a robust representation of the background.



Block diagram Proposed Methodologies

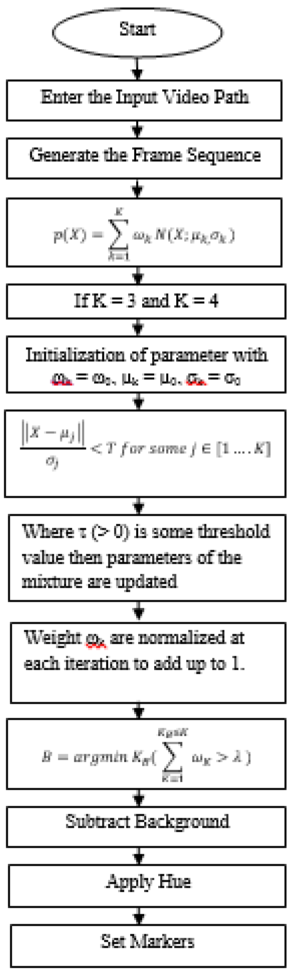
The work starts with the generation of frame sequences from the input video. Then the Probability density functions are calculated for a Gaussian Mixture comprising K component densities. MFCCs are obtained as follows [17] [41] First, the parameters are initialized with $\omega_k = \omega_0$, $\mu_k = \mu_0$ and $\sigma_k = \sigma_0$. If there is a match, i.e. then the parameter mixtures are updated as per the mentioned equations follow up by normalizing the weights at each iteration to add up to 1. A threshold λ is applied to the cumulative sum of weights to find the set {1...B} of Gaussians modeling the background. Intuitively, Gaussians with the highest probability of occurrence, wk, and lowest variability in the distribution, measured by σ_k, indicating a representative mode, are the most likely to model the background.

Parameters for Simulation

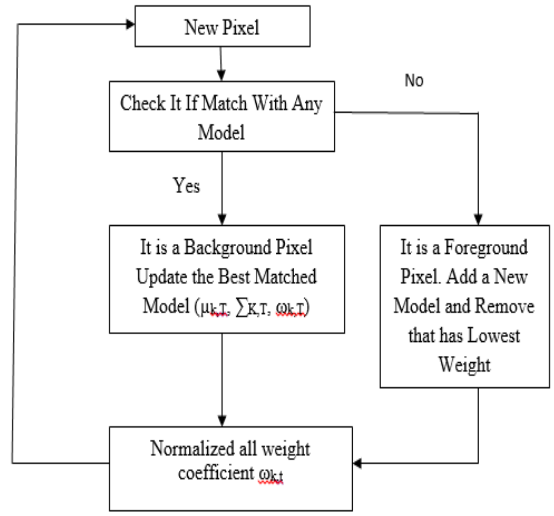
The parameters that have been used in simulation are mentioned and briefly discussed below;

1) Number of Gaussian Densities (K): It represents the number of Gaussian densities used that are used to compute the PDF. Calculations have been done for K=3 and K=4.
2) Background Threshold (λ): A threshold λ is

3) applied to the cumulative sum of weights to find the set {1...B} of Gaussians modelling the background.
4) Covariance (σ): Covariance matrix which is used in calculation of initial pdf.
5) Component Threshold: Normally taken as 10.



Flow diagram of proposed methodologies



Flow chart of Updating MOG's Model

SIMULATION RESULTS

Car Park Video

1) 520 frame Video.
2) 10 fps.
3) Background: Stable.
4) Illumination Change: Partial.
5) Objects to track: Multiple.

The video consists of multiple objects that are required to be tracked. The system efficiently tracks both the moving car and the pedestrian. It locks on to moving man once the car is stationary, and that the multiple objects have been tracked successfully. The algorithm has seamlessly detected even multiple objects as it can be seen from various images where after subtraction and morphological filtering correct markers have been implanted.

Input frame from video that consist of multiple objects like moving car, stationary car and moving man also. This frame we can see in the Fig.1

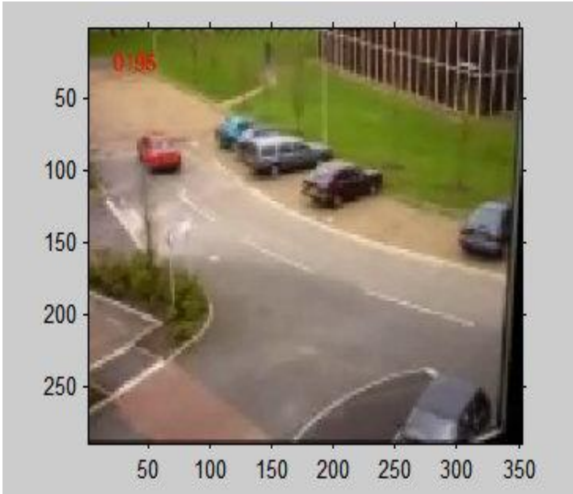


Fig 1 Input Frame

To track the multiple objects we have to extract the best background of this input video and this extracted image is extracted best background image. As we can see from the Fig. False detection is more prominently visible in the initial learning stage that should be removing after using some applications. When some false detection occurred in tracked object image after subtraction frame then we updated the mixture parameter, and the object is traced successfully with few false detection being removed by filtering. As we can see in fig 2.

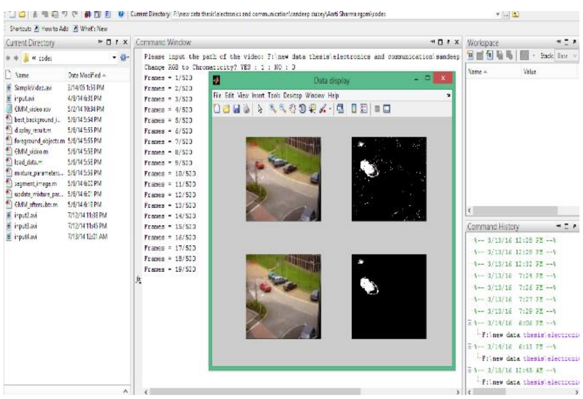


Figure 2 Extracted Best Background Image.

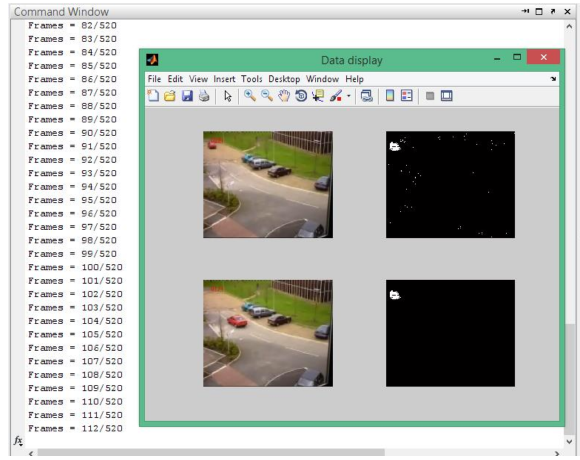


Fig. 3 Initial learning phase after filtering.

After filtering of image frame of initial learning phase apply hue to detected area for tracking the object. As we can see in the results Fig. 4.

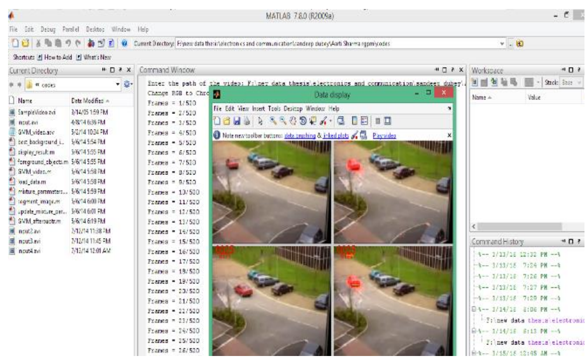


Fig. 4after applying Hue to detected area.

Applying Hue to detected area we have to show the object so we marked the object and traced the object successfully. As we can see from the Fig. 6.

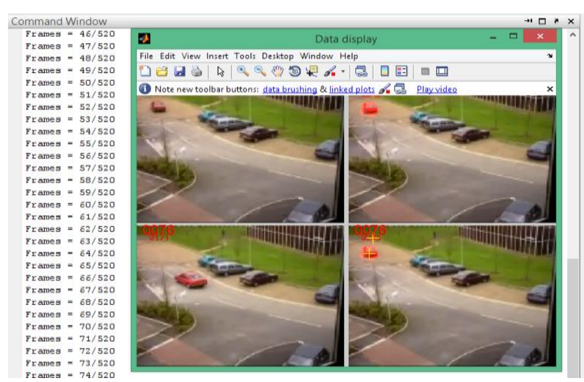


Fig. 6Object Marked and Tracked.

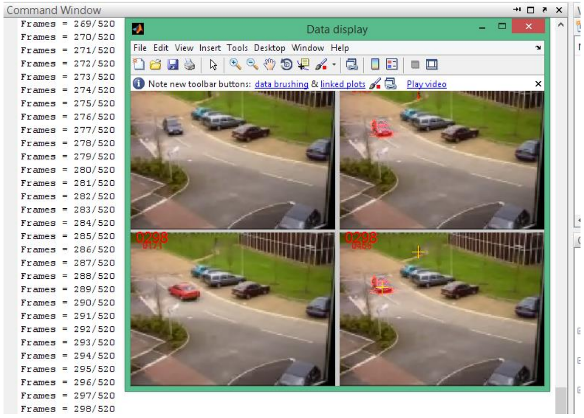


Fig. 7 Multiple Object Detection (Man and car both moving).

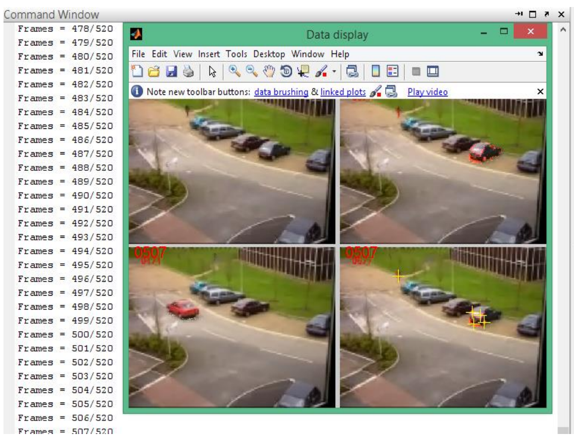


Fig. 8 Car Stationary and man moving.

The video consists of multiple objects that are required to be tracked. The system efficiently tracks both the moving car and the pedestrian. It locks on to moving man once the car is stationary. However the initial learning phase was slightly slower than previous videos owing to the initial visibility in this video is very poor as the illumination change is significant and the camera is at a significant distance away from the object.

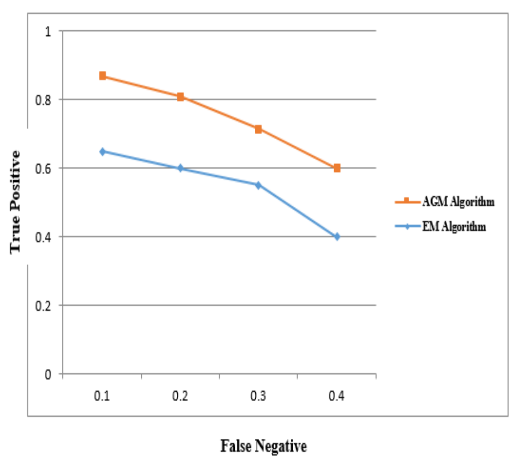
Comparison of Past and Present Work

Sr. No.	Parameter	Past work	Present work
1	Modeling method	GMM, EM Algoritm	AGM, Pixel Based Process
2	Change of illumination and lighting adaption	Not adapted	Adapted
3	Motion sensing and tracking	Individually	Simultaneously
4	Comprising component densities	K = 2	K = 4
5	α , RHO	α = 0.001 RHO = 0.001	α = 0.002 RHO = 0.002
6	Deviation Threshold	Not fixed	Fixed
7	Initial Variance	8.97588967631258	7.99892827140397
8	Background Threshold	0.81895021796423	0.95619778887127
9	Component Threshold	7	10
10	Accuracy	Poor	Good
11	Initial Mix prop	0.00985684713287	0.00868042059550
12	Average Processing Time	10 ms	8 ms
13	Economy	Higher Cost	Low Cost
14	Power	More	Less
15	Average True Positive Recognition	75.5%	85%

Evaluation Value True Positive and False Negative Recognition

Parameter	True Positive Rate Value		False Negative Rate Value	
Algorithm / Variance	EM Algorithm	AGM Algorithm	EM Algorithm	AGM Algorithm
0.012	0.65	0.85	0.12	0.12
0.012	0.6	0.81	0.20	0.20
0.012	0.57	0.70	0.30	0.30
0.012	0.40	0.60	0.40	0.40

Using the value of evaluation of performance of proposed algorithm (AGM Algorithm) and past work algorithm (EM Algorithm) we conclude that the mean of true positive is 85% with variance 0.012 and the mean of false negative rate is 0.12. We can also see from the graph which has plotted between true positive recognition and false negative Recognition rate.



Line Chart of Evaluation of Performance of Algorithm

Conclusion

This paper has presented a detailed account on the state of the art in the field of Motion Detection through Computer Vision. The work discussed all the technologies like Optical flow, Gaussian average etc. and the mathematical concepts involved in the algorithms. The paper discussed at length the advantages using Gaussian Mixture models and presented the use of Adaptive GMM as an enhanced tool for motion sensing. The results showed the effectiveness of AGMM in detection of motion in videos with varying light intensities and poor visibilities. The work showed satisfactory performance in terms of its detection capabilities and learning rate performance.

References

1) Kalal Z, Miko lajczyk K, Matas, J., "Tracking-Learning-Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume: 34, Issue: 7 Page(s):1409 – 1422) July 2012.
2) Berthold K. Horn ; Brian G. Schunck, "Determining Optical Flow", Proc. SPIE 028 Conference on Techniques and Applications of Image Understanding, Washington, April 21, 1981.
3) Antonio Fernández, Caballero A, José Carlos Castillo a, Javier Martínez-Cantos c,, "Optical flow or image subtraction in human detection from infrared camera on mobile robot"s, Robotics and Autonomous Systems , Pp 1273–1281, ScienceDirect (2010).
4) Jaesik Choi, "Realtime On-Road Vehicle Detection with Optical Flows andHaar-Like Feature Detectors", F. T. Luk, ed., SPIE-The International Society for Optical Engg. Proceedings 2563, SPIE, Washington, DC, pp. 314-325, 1995.
5) S. S. Beauchemin , J. L. Barron, "The computation of optical flow", Journal of ACM Computing Surveys (CSUR) Volume 27 Issue 3, , Pages 433-466, Sept. 1995.
6) O. Barnich and M. Van Droogen broeck. Vi Be: A Universal BackgroundSubtraction Algorithm for Video Sequences. IEEE Transactions on Image Processing, 20(6):1709–1724, June 2011.
7) Ahmed Elgammal, David Harwood, and Larry Davis. Non-parametric model forbackground subtraction. In FRAME-RATE WORKSHOP, IEEE, pages 751–767, 2000.
8) K Toyama, J Krumm, B Brumitt, and B Meyers. Wallflower: principles andpractice of background maintenance. Proceedings of the Seventh IEEE International Conference on Computer Vision, 1(c):255–261, 1999.
9) L. Maddalena and A. Petrosino. A Self-Organizing Approach to BackgroundSubtraction for Visual Surveillance Applications. IEEE Transactions on Image Processing, 17(7):1168–1177, 2008.
10) Ismail Haritao glu, David Harwood, and Larry S. Davis. W4: Real-timesurveillance of people and their activities. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22:809–830, 2000.

11) DA Migliore, M Matteucci, M Naccari – "A revaluation of frame difference in fast and robust motion detection" Proceedings of the 4th ACM, dl.acm.org.- 2006.
12) Chris Stauffer, W. Eric, and W. Eric L. Grimson. Learning patterns of activityusing real-time Tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22:747–757, 2000.
13) Alan J. Lipton, Hironobu Fuji yoshi, and Raju S. Patil. Moving targetclassification and tracking from real-time video. In Proceedings of the 4th IEEE Workshop on Applications of Computer.
14) Douglas Reynolds ,Gaussian Mixture Models, MIT Lincoln Laboratory,1998.
15) Gray, R. Vector Quantization. IEEE ASSP Magazine pages: 4–29 (1984).
16) Reynolds, D.A. A Gaussian Mixture Modelling Approach to Text-Independent Speaker Identification. PhD thesis, Georgia Institute of Technology (1992).
17) Reynolds, D.A., Rose, R.C.: Robust Text-Independent Speaker Identification using Gaussian Mixture Speaker Models. IEEE Transactions on Acoustics, Speech, and Signal Processing 3(1) 72–83 (1995).
18) C. Stauffer and W. E. L. Grimson, "Adaptive back ground mixture models for real-time tracking," in Proc. of IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 246–252, 1999.
19) N. McFarlane and C. Schofield, "Segmentation and tracking of piglets in images," Machine Vision and Applications, 8(3), pp. 187-193, 1995.
20) I. Haritao glu, D. Harwood, and L. Davis, W4, "Real-time surveillance of people and their activities," IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 22, no. 8, 2000.
21) R.Collins, A.Lipton, and T.Kanade, "A system for video surveillance and monitoring," in Proceedings of American Nuclear Society (ANS) Eighth International Topical Meeting on Robotics and Remote Systems, pp. 25-29, April 1999.
22) A. Elgammal, R. Duraiswami, D. Harwood, and L. Davis, "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," in Proceedings of the IEEE, pp. 1151-1163, July 2002}.
23) N. Friedman and S. Russell., Image segmentation in video sequences: A probabilistic approach," in Proceedings of Thirteenth Conference on Uncertainty in Artificial Intelligence, pp. 175-181, 1997.
24) F. Porikli and O. Tuzel, "Human body tracking by adaptive background models and mean-shift analysis," in IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, March 2003.
25) C. Wren, A. Azar bayejani, "T. Darrell, and A. Pent land: Real-time tracking of the human body," IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 19, pp. 780-785, July 1997.
26) K. Toyama, J. Krumm, B. Brumitt, and B. Meyers, "Principles and practice of background maintenance," in Proceedings of IEEE International Conference on Computer Vision (ICCV), pp. 255-261, 1999.
27) M. M. A. J. Manoj S. Nagmode, Ashok M. Sapkal, "A novel approach to detect and track moving object using Partitioning and Normalized Cross Correlation," ICGST-GVIP vol.9, 2009.
28) Robert T. Collins, Alan J. Lipton, Takeo Kanade, "A System for Video Surveillance and Monitoring," The Robotics Institute, Carnegie Mellon University, Pittsburgh PA, 2012.
29) P. Kaew TraKul Pong, R. Bowden, "An Improved Adaptive Background Mixture Model for Real-time Tracking with Shadow Detection 2 Background Modelling," Online (Weston, Conn.), vol. 1, p. 1, 2001.
30) R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting moving objects, ghosts, and shadows in video streams," IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 25, October 2003.
31) S. J. McKenna, S. Jabri, Z. Duric, and A. Rosenfeld, "Tracking groups of people," Computer Vision and Image Understanding (CVIU), no. 80, pp. 42-56, 2000.
32) Y. Yang and M. Levine, "The background primal sketch an approach for tracking moving objects," Machine Vision and Applications, vol. 5, pp. 17-34, 1992.
33) S.-C. S. Cheung and C. Kamath, "Robust techniques for background subtraction in urban tracking video," in Proceedings of SPIE, Visual Communications and Image Processing 2004, vol. 5308, pp. 881-892, January 2004.

34) P. Kaew TraKul Pong, R. Bowden, "An Improved Adaptive Background Mixture Model for Real-time Tracking with Shadow Detection 2 Background Modelling," Online (Weston, Conn.), vol. 1, p. 1, 2001.

35) D. H. Thanarat Horprasert, Larry S. Davis, "A Statistical Approach for Real-time Robust Background Subtraction and Shadow Detection," presented at the IEEE ICCV'99FRAME-RATE WORKSHOP, 1999.

36) E. Hayman and J. O. Eklundh, "Statistical background subtraction for a mobile observer, "in Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on, pp. 67-74 vol.1, 2003.

37) O. Javed, K. Shafique, and M. Shah, "A Hierarchical Approach to Robust Background Subtraction using Colour and Gradient Information," presented at the Proceedings of the Workshop on Motion and Video Computing, 2002.

38) Ahmed Elgammal, Ramani Duraiswami," Background and Foreground Modelling Using Non parametric Kernel Density Estimation for Visual Surveillance," Proceedings Of The IEEE, Vol. 90, No. 7, July 2002.

39) B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," presented at the Proceedings of the 7th international joint conference on Artificial intelligence - Volume 2, Vancouver, BC, Canada, 1981.

40) R. Szeliski. Computer Vision: Algorithms and Applications (2010).