A Review Analysis of Different Object Detection and Subtraction Mechanism

[1]Rekha Tiwari (M-Tech Scholar, JNCT, Rewa, India)
rekhatiwa@gmail.com
[2]Sumit Dubey (Assistant Professor ECE Department, JNCT, Rewa, India)
[3]Sneha Singh (HOD, ECE Department, JNCT, Rewa, India)
Singhsneha1888@gmail.com

Abstract
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.

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.

Applications of Motion Detection

Motion Detection as already stated has found its applications in almost all forms of life where an active monitoring system is required. The applications range from common household monitoring systems to advanced state of the art military Intelligence systems. Some of the Premier applications of the systems have been listed below [1].

- Motion-based recognition including human identification based on gait, gestures, automatic object detection, etc.
- Automated surveillance: monitoring a scene to detect suspicious activities or unlikely events. Intrusion detection, Burglar alarm systems, anomalous animal behavior in parks etc.
- Video indexing: automatic annotation and retrieval of videos in multimedia databases.
- Human-computer interaction: gesture recognition, eye gaze tracking for data input to computers, etc.
- Traffic monitoring: Gathering of traffic statistics for controlling and redirecting traffic based on inputs.
- Vehicle navigation: video-based path planning and obstacle avoidance capabilities.

Before going into details of our proposed methods, we first introduce the concept is used both in the following and through the thesis to make a clear understanding.

1. Frame I is used to denote one frame from a video sequence and it is used to denote the

frame at time t or the t_{th} frame in the sequence.

2. Pixel value at location (x; y) is represented by I(x,y), which a vector of length 1 if the image is a gray or binary image, a vector of length 3 if the image is a colour image.
3. Background image B is used frequently referring to the image from a video sequence with no moving objects. The background image can be fixed or can be updated with time denoted as B_t.
4. Absolute difference image D. We will compare a frame It with a reference image to get an absolute difference image D_t. The reference image could be the background image B_t, then $D_t = |I_t - B_t|$ or the reference image could be the frame before I_t, then $D_t = |I_t - I_{t-1}|$ and it is called consecutive difference image alternatively.
5. Foreground or foreground region denotes the region where motion occurs, which is also the target of motion detection process. In this thesis, our interested foreground regions are those regions occupied by moving persons.
6. Background or background region refers to the image region which is static comparing with foreground, for example the room one person is walking in. Foreground and background together form the whole image at the time of motion detection.

Motion Detection Techniques

Various Algorithms have been proposed for Motion sensing, detection and Tracking purposes. Ranging from basic frame differencing to more advanced algorithms like TLD by Zdenek Kalal. The performances vary depending on the types of backgrounds, frame rates, learning rates etc. Based upon these metrics various mechanisms have been discussed in this section. Motion detection techniques are broadly classified in to two main categories;

1) Region based Algorithms
2) Pixel Based Algorithms

Region based algorithms due to their spatial dependencies of neighboring color pixels however the latter are based on binary differences by employing local or pixel based model of Intensity. Being simple, they have their application in real time solutions as well.

Problems and Issues

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

Background Subtraction

The background subtraction [6-10] is the most popular and common approach for motion detection. The idea is to subtract the current image from a reference background image, which is updated during a period of time. It works well only in the presence of stationary cameras. The subtraction leaves only non-stationary or new objects, which include entire silhouette region of an object.

These approaches are simple and computationally affordable for real-time systems, but are extremely sensitive to dynamic scene changes from lightning and extraneous event etc. Therefore it is highly dependent on a good background maintenance model. Here in this chapter we have simulated different background subtraction techniques available in the literature for motion segmentation of object. Background subtraction detects moving regions in an image by taking the difference between the current image and the reference background image captured from a static background during a period of time. The subtraction leaves only non-stationary or new objects, which include entire silhouette region of an object. The problem with background subtraction [8-9] is to automatically update the background from the incoming video frame and it should be able to overcome the following problems;

1. Motion in the background:
   Non-stationary background regions, such as branches and leaves of trees, a flagwaving in the wind, or flowing water, should be identified as part of the background.
2. Illumination changes:
   The background model should be able to adapt, to gradual changes in illumination over a period of time.
3. Memory:
   The background module should not use much resource, in terms of computing power and memory.
4. Shadows:
   Shadows cast by moving object should be identified as part of the background and not foreground.
5. Camouflage:

Moving object should be detected even if pixel characteristics are similar to those of the background.

6. Bootstrapping:
   The background model should be able to maintain background even in the absence of training background (absence of foreground object).

The principal background subtraction techniques are;

1. Using frame differencing
2. Selectivity
3. Running Gaussian average
4. Background mixture models

Frame Differencing

Frame differencing [12] is a pixel-wise differencing between two or three consecutive frames in an image sequence to detect regions corresponding to moving object such as human and vehicles. The threshold function determines change and it depends on the speed of object motion. It’s hard to maintain the quality of segmentation, if the speed of the object changes significantly. Frame differencing is very adaptive to dynamic environments, but very often holes are developed inside moving entities. We have secured many results when we apply simple frame differencing to input video frame. Such as input video frame for simple frame differencing, Foreground mask obtained through simple frame differencing, Frame difference results with threshold set at high and low.

Input Video Frame for simple frame differencing.

Foreground mask obtained through simple frame differencing

Frame difference is normally calculated as;

Frame difference=$|\,frame_i - frame_{i-1}| > Th$

Here the estimated background is just the previous frame. It evidently works only in particular conditions of objects speed and frame rate. However, results are very sensitive to the threshold Th.

Frame difference results with Thresholds set at high and low.

Background Subtraction Using Gaussian Mixtures

As computer vision begins to address the visual interpretation of action applications such as surveillance and monitoring are becoming more relevant. Similarly, recent work in intelligent environments and perceptual user interfaces involve vision systems which interpret the pose or gesture of users in a known, indoor environment. In all of these situations the first fundamental problem encountered is the extraction of the image region corresponding to the object or persons in the room. Previous attempts at segmenting object from a known background have taken one of the three approaches mentioned previously. Most common is some form of background subtraction. For example, Grimson et al. uses statistical texture properties of the background observed over extended period of time to construct a model of the background, and use this model to decide which pixels inan input image do not fall into the background class.

The fundamental assumption of the algorithm is that the background is static in all respects: geometry, reflectance and illumination [13].

The second class of approach is based upon image motion only presuming that the background is stationary or at most slowly varying, but that the object is moving.

In these methods no detailed model of the background is required. Of course, these methods are only appropriate for the direct interpretation of motion; if the object stops moving, no signal remains to be processed. This method also requires constant or slowly varying geometry, reflectance and illumination.

Gaussian Mixture Models

A Gaussian Mixture Model (GMM) is a parametric probability density function represented as a weighted sum of Gaussian component densities. GMMs are commonly used as a parametric model of the probability distribution of continuous measurements or features in a biometric system, such as vocal-tract related spectral features in a speaker recognition system. GMM parameters are estimated from training data using the iterative Expectation-Maximization (EM) algorithm or Maximum A. Posteriori (MAP) estimation from a well-trained prior model.

A Gaussian mixture model is a weighted sum of M component Gaussian densities as given by.

$$p\left(\frac{X}{\lambda}\right) = \sum_{i=0}^{M} \omega_i \; g\left(\frac{X}{\mu_i}\right), \Sigma_i$$

where x is a D-dimensional continuous-valued data vector (i.e. measurement or features), ω_i, i = 1, . . . ,M are the mixture weights, and $g(x|\mu_i,\sum i)$, i = 1, . . . ,M are the component Gaussian densities.

Each component density is a D-variate Gaussian function of the form [13].

$$g(x|\mu_i,\textstyle\sum i) = \frac{1}{2\pi^{D/2}|\sum i|^{1/2}} \exp\{-1/2(X-\mu_i)'\textstyle\sum -1(X-\mu_i)\}$$

With mean vector μ_i and covariance matrix $\sum i$. The mixture weights satisfy the constraint that $\sum_{i=1}^{M} \omega_i = 1$.The complete Gaussian mixture model is parameterized by the mean vectors, covariance matrices and mixture weights from all component densities. These parameters are collectively represented by the notation.

$\lambda = \{\omega_i, \mu_i, \Sigma_i\}$ i=1,2,3……….,M

There are several variants on the GMM. The covariance matrices $\sum i$, can be full rank or constrained to be diagonal. Additionally, parameters can be shared or tied among the Gaussian components such as having a common covariance matrix for all components. The choice of model configuration (number of components full or diagonal covariance matrices and parameter tying) is often determined by the amount of data available for estimating the GMM parameters and how the GMM is used in a particular biometric application.

It is also important to note that because the component Gaussian is acting together to model the overall feature densities, full covariance matrices are not necessary even if the features are not statistically independent. The linear combination of diagonal covariance basis Gaussians is capable of modeling the correlations between feature vector elements. The effect of using a set of M full covariance matrix Gaussians can be equally obtained by using a larger set of diagonal covariance Gaussians.

GMMs are often used in biometric systems most notably in speaker recognition systems due to their capability of representing a large class of sample distributions. One of the powerful attributes of the GMM is its ability to form smooth approximations to arbitrarily shaped densities. The classical uni-modal Gaussian model represents feature distributions bya position (mean vector) and an elliptic shape (covariance matrix) and a vector quantizer (VQ) or nearest neighbor model represents a distribution by a discrete set of characteristic templates [13]. A GMM acts as a hybrid between these two models by using a discrete set of Gaussian functions, each with their own mean and covariance matrix, to allow a better modeling capability. Compares the densities obtained using a unimodal Gaussian model, a GMM and a VQ model.

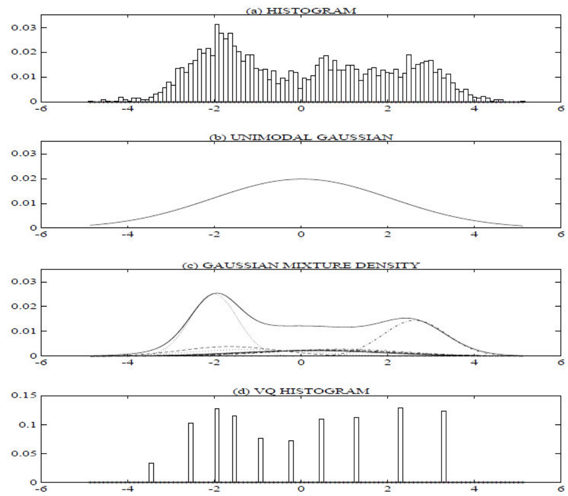
Plot (a) shows the histogram of a single feature from a speaker recognition system (a single cepstral value from a 25 second utterance by a male speaker).

Plot (b) shows a uni-modal Gaussian model of this feature distribution.

Plot (c) shows a GMM and its ten underlying component densities.

Plot (d) shows a histogram of the data assigned to the VQ centroid locations of 10 element code book. The GMM not only provides a smooth overall distribution fit, its components also clearly detail the multi-modal nature of the density.

The use of a GMM for representing feature distributions in a biometric system may also be motivated by the intuitive notion that the individual component densities may model some underlying set of hidden classes. For example, in speaker recognition, it is reasonable to assume the acoustic space of spectral related features corresponding to a speaker's broad phonetic events, such as vowels, nasals or fricatives. As we can see in the use of GMM in speaker recognition biometric system.



Use of GMM in Speaker recognition Biometric system.

These acoustic classes reflect some general speaker dependent vocal tract configurations that are useful for characterizing speaker identity. The spectral shape of the i^{th} acoustic class can in turn be represented by the mean μ_i of the i^{th} component density and variations of the average spectral shape can be represented by the covariance matrix i. Because all the features used to train the GMM are unlabeled, the acoustic classes are hidden in that the class of an observation is unknown. A GMM can also be viewed as a single-state HMM with a Gaussian mixture observation density, or an ergodic Gaussian observation HMM with fixed, equal transition probabilities. Assuming independent feature vectors, the observation density of feature vectors drawn from these hidden acoustics classes is a Gaussian mixture [15-16].

Adaptive Mixture of Gaussian

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.

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

Where ω_k are the weights and $N(x; \mu_k, \sigma_k)$ is the normal density of mean μ_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.
First, the parameters are initialized with $\omega_k = \omega_0$, $\mu_k = \mu_0$ and $\sigma_k = \sigma_0$. If there is a match.

$$\frac{\|X - \mu_j\|}{\sigma_j} < T \text{ for some } j \in [1 \ldots . K]$$

Where τ (> 0) is some threshold value, then the parameters of the mixture are updated as follows.

$$\omega_k(t) = (1-\alpha)\omega_k(t-1) + \alpha M_k(t)$$

$$\mu_k(t) = (1-\beta)\mu_k(t-1)$$

$$\sigma_k^2 t^2 = (1-\beta)\sigma_k^2(t-1) + \beta \|X - \mu_k(t)\|2$$

Where $M_k(t)$ is equal to 1 for the matching component j and 0 otherwise. If there is no match, the component with the lowest weight ω_k is re-initialized with $\omega_k = \omega_0$, $\mu_k = x$ and $\sigma_k = \sigma_0$. The learning rate α is constant and β is defined as;

$$\beta = \alpha N(\mathbf{x}; \mu_k, \sigma_k).$$

Finally, the weights wk are normalized at each iteration to add up to 1. Stauffer and Grimson proposed to sort the Gaussians by decreasing weight-to-standard-deviation ratio ω_k/σ_k, to represent the background. A threshold λ is applied to the cumulative sum of weights to find the set {1...B} of Gaussians modeling the background, defined as.

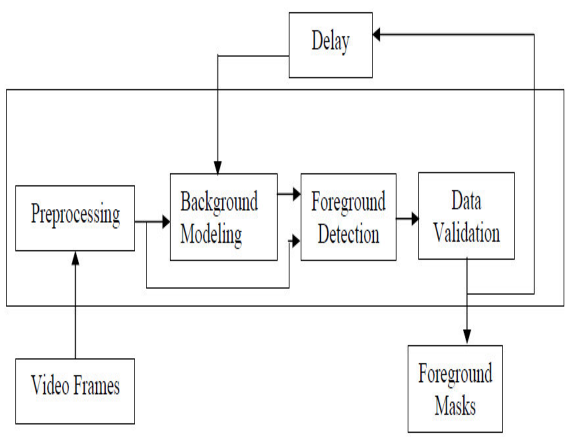
$$B = argmin\ K_B\left(\sum_{K=1}^{K_B \le K} \omega_K > \lambda\right)$$

Intuitively, Gaussians with the highest probability of occurrence, ω_k and lowest variability in the distribution measured by σ_k, indicating a representative mode are the most likely to model the background.

Generic Background Subtraction Algorithm

Even though there exist a myriad of background subtraction algorithms in the literature [16] most of them follow a simple flow diagram shown in figure. The four major step in a background subtraction algorithm are preprocessing, background Modeling, foreground detection and data validation. Preprocessing consists of a collection of simple image processing tasks that change the raw input video into a format that can be processed by subsequent steps.

Background modeling uses the new video frame to calculate and update a background model. This background model provides a statistical description of the entire background scene. Foreground detection then identifies pixels in the video frame that cannot be adequately explained by the background model and outputs them as a binary candidate foreground mask. Finally, data validation examines the candidate mask, eliminates those pixels that do not correspond to actual moving objects and outputs the final foreground mask. Domain knowledge and computationally-intensive vision algorithms are often used in data validation. Real-time processing is still feasible as these sophisticated algorithms are applied only on the small number of candidate foreground pixels. Many different approaches have been proposed for each of the four processing steps. Some of the representative ones in the following subsections have been reviewed.



Flow diagram of generic background subtraction algorithm.

Past Work

Human motion detection is a fundamental research area in computer vision. Many vision applications require segmenting the motion region out of the scene, which is usually called motion detection. Motion detection is an important part of many computer vision tasks like human tracking, pose estimation and face recognition. This chapter presents an overview of the state of the art in the field of video based motion detection. Since motion is a temporal event, most motion detection methods use temporal information from adjacent images or a much longer image sequence [20-23]. The most popular motion detection method is frame subtraction i.e. a current frame is compared pixel-wise with a reference image. If a pixel value is above a preset value, it is assumed to be brought by motion. Using a static camera to observe a scene is quite common in a smart room application or a surveillance system [20] [24] [25]. The static scene is often referred to as background and the moving object is referred to as foreground.

Many motion detection methods have been extensively investigated [23] [26] gave a good discussion of the research methods. One direct method is to use temporal difference. The absolute difference at each pixel between two or three consecutive frames is calculated and a threshold is applied to get the difference image.

Background subtraction uses only a single frame of background as the model. An image subtraction between the input frame and the model followed by thresholding is implemented to determine foreground pixels.

$$I_{foreground} = |I_{input} - I_{background}| > T$$

Where $I_{foreground}$ is the foreground objects image, I_{input} is the input image, $I_{background}$ is the background image, and T is the difference threshold. This simple model only works in the ideal case, where the background is fixed and is not affected by lighting changes or vibration. In practice, the subtraction image is really noisy.

In [28] a three frame difference algorithm was used. Their three-frame differencing rule suggested that a pixel was legitimately moving if its intensity had changed significantly between both the current image and the last frame, and the current image and the next-to-last frame. This method was simple to implement and it could adapt fast to background changes, but it was not so effective to get the whole region of the moving object due to the following reason: since absolute difference was used, the difference image may include both the pixels which were previously background but now covered by foreground and pixels that were previously foreground but became uncovered background. On the other hand, if the motion was not big between frames, the inner part of the moving object cannot be detected.

In [29] a connected component analysis was used to cluster the difference image into motion regions to facilitate further processing. In [29] besides using motion information from two consecutive frames (frame difference as in the paper), they also constructed and maintained an up-to-date background model from the video sequence and compared each frame with the background (background difference as in the paper). Frame difference and background difference were combined together to detect motion more precisely.

Selecting color model is also important to reduce the effect of lighting changes. Color spaces such as YCbCr and HSV separate color from intensity and makes the algorithm more robust to changing intensity (i.e. lighting changes due to the time of day) or simpler to detect shadows or to model the color for tracking. Because the background subtraction method is simple, its application is limited to the indoor environment, where the background appearance is assumed to be consistent overtime. For outdoor environments, a single model is not sufficient to cope with variations in lighting of the background. Multi-modal approaches [29] have been applied to solve the practical problems such as time varying backgrounds or lighting variations. The multi-modal solution stores numerous models of the background for each pixel, under the probabilistic model.

$$P(I_{xy} \in B) = \sum_{i=1:N} \omega_i P(I_{xy} \in M_i)$$

Using statistical background model is a more popular method to do motion detection [30]. A simple background model can be the average image over some training period. Motion can be detected by thresholding the difference between the mean background model and the current image. Instead of using a threshold, in [31] the pixel mean and variance of the R, G and B channel were stored for each pixel as background model and were updated recursively. A current pixel was compared to the model, if in either channel the distance between current pixel and the mean value of background model was greater than 3 times the standard deviation, the pixel was set to foreground. Otherwise it was set to background. Some researchers claim that the median value was more robust than the mean value [32]. Cucchiara et al [30] modeled the background using median function; they report that the median function had proven effective while at the same time of less computational cost than using complex statistics like mixed Gaussian model. Cheung and Kamath [33] also reported similar results.

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