Video Surveillance Systems – A Survey

C. Lakshmi Devasena [1], R. Revathí [2], M. Hemalatha [3]

**[1] Department of Software Systems, Karpagam University,
Coimbatore, Tamilnadu, India**

**[2] Research Scholar, Karpagam University,
Coimbatore, Tamilnadu, India**

**[3] Department of Software Systems, Karpagam University,
Coimbatore, Tamilnadu, India**

Abstract

Video surveillance is increasing significance approach as organizations seek to safe guard physical and capital assets. At the same time, the necessity to observe more people, places, and things coupled with a desire to pull out more useful information from video data is motivating new demands for scalability, capabilities, and capacity. These demands are exceeding the facilities of traditional analog video surveillance approaches. Providentially, digital video surveillance solutions derived from different data mining techniques are providing new ways of collecting, analyzing, and recording colossal amounts of video data. This paper addresses some of the approaches for video surveillance systems.

Keywords: *Automatic video surveillance, Object Tracking, Multimedia surveillance system, Real Time Video surveillance.*

1. Introduction

Human motion analysis helps in solving problems in indoor surveillance applications. K. Srinivasan et al [4] presented an attempt to give an idea of human body tracking in surveillance area of monocular video sequences. They have discussed various kinds of background modeling techniques like Background Subtraction Method, Adaptive background subtraction method, Adaptive Gause Mixture Method, 2D and 3D human body tracking methods etc. Human body modeling identifies the body positions and activities in video sequences. The model proposed by them is shown in Figure 1.

The frame work of the system starts with the acquiring of video images by means of camera and pre-processing has to be done on them for enhancing the quality of frames in the sequences. The video frames have a lot of noise due to camera, illumination and reflections etc. This can be removed and quality of images can be enhanced with the help of preprocessing stages. The suitable steps should be carried out in this stage.

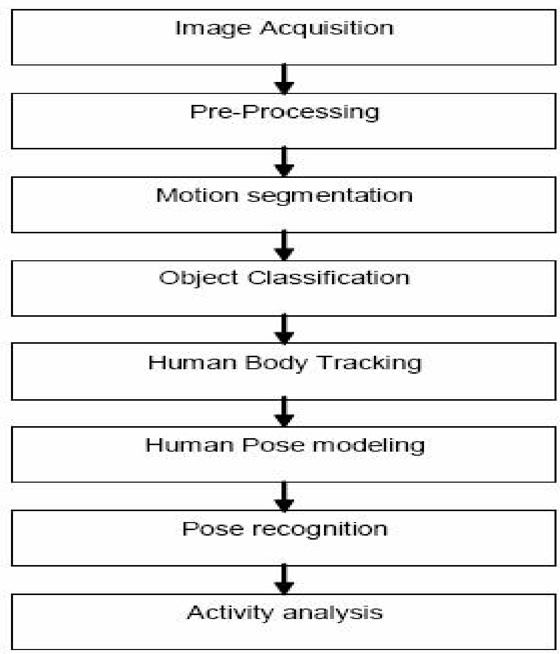

Figure 1: System Model of the Video Surveillance System

The next stage is motion segmentation which separates foreground images from background images and it is followed by Object classification, Tracking and Human pose modeling. At the end, the activity analysis will be processed.

Massimo Piccardi [6] reviewed about eight background subtraction techniques used for object tracking in video surveillance ranging from simple approaches, used for maximizing speed and restraining the memory requirements, to more complicated approaches, used for accomplishing the highest possible accuracy under any potential circumstances. All approaches intended for real-time performance. The techniques reviewed are: Running

Gaussian average, Temporal median filter, Mixture of Gaussians, Kernel density estimation (KDE), Sequential KD approximation, Cooccurence of image variations and Eigen backgrounds technique.

Amongst the methods reviewed, simple methods such as the Gaussian average or the median filter offer acceptable accuracy while achieving a h igh frame rate and having limited memory requirements. Methods such as Mixture of Gaussians and KDE prove very good model accuracy. KDE has a high memory requirement (in the order of a 100 frames) which might prevent easy implementation on low-memory devices. SKDA is an approximation of KDE which proves almost as accurate, but mitigates the memory requirement by an order of magnitude and has lower time complexity. Methods such as the Cooccurence of image variations and the Eigen backgrounds explicitly address spatial correlation. They both offer good accuracy against reasonable time and memory complexity.

2. Video Surveillance for Real-Time Applications

Muller-Schneiders et al [7] presented a f ull fledged introduction to the real time video surveillance system which has robustness as the major design goal. A robust surveillance system should specifically aimed for a low volume of false positives because surveillance guards might get deviated by too many alarms caused by, e.g., moving rain , trees, , varying illumination conditions or small camera motion. Since a missed security related event could cause a powerful threat for an installation site, the previously mentioned criterion is not enough for designing a robust system and thus false negatives should simultaneously be achieved. Due to the requirement that the false negative rate should be equal to zero, the surveillance system should cope with varying illumination conditions, occlusion situations and low contrast. Apart from presenting the building blocks of the video surveillance system, the measures taken to achieve robustness is illustrated. Since their system is based on algorithms for video motion detection. To measure the performance of the system, quality measures are calculated for various PETS.

Real-time detection of moving objects is vital for video surveillance. Nan Lu et al [8] has proposed a n ovel real time motion detection algorithm which integrates the temporal differencing method, double background filtering (DBF) method, optical flow method and morphological processing methods to obtain fine performance. The temporal differencing is used to reveal initial coarse motion areas for the optical flow calculation to arrive real-time and accurate object motion detection. The double background filtering method is used to achieve and keep a steady background image to manage with differences on environmental changing conditions and is used to eradicate the background interference information and detach the moving object from it. The morphological processing methods are applied and combined with the DBF to get enhanced results. The most gorgeous advantage of this algorithm is that the algorithm does not necessitate learning the background model from hundreds of images. It can handle quick image dissimilarities without former knowledge about the object shape and size. The algorithm has high ability of anti-interference and maintains high accurate rate detection at the same time. It also demands a reduced amount of computation time than other methods for the real-time surveillance. The efficacy of the anticipated algorithm for motion detection is established in a simulation environment and the evaluation results are reported as fine.

To solve the p roblems like excessive storage space required to store the video, time consumption to record and view the video, C hing-Kai Huang and Tsuhan Chen [1] proposed a method by recording only video that has important information, i.e., video that has motion in the scene. This can be achieved with a d igital video camera and a DSP algorithm that detects motion. They have reported a v ideo surveillance system that was developed based on TI DSP ‘C54 which is cheap, small and requires less power. And their algorithm called block-based MR-SAD (Mean Reduced – Sum Average Difference) method is used to robustly distinguish the motion from lighting changes by removing the mean from the frame difference signal. Their algorithm decomposes the image into small blocks, which optimistically separate objects with different reflectivity into different blocks. Then, for each block of the frame difference, the mean is calculated and it is subtracted from the frame difference. After that, absolute value of all pixels is summed up. The system which uses a built-in C54 to trigger the recording mechanism can cut down the cost of the storage space significantly.

Lun Zhang et al [2] described an appearance-based method to accomplish real-time and vigorous objects classification in varied camera viewing angles. A new descriptor called, the Multi-block Local Binary Pattern (MB-LBP), is used to capture the large-scale structures in object appearances. Based on MB-LBP features, an adaBoost algorithm is brought in to select a subset of discriminative features as well as construct the strong two-class classifier. To deal with the non-metric feature value of MBLBP features, a multi-branch regression tree is developed as the weak classifiers of the boosting. At last, the Error Correcting Output Code (ECOC) is established to achieve vigorous multi-class classification performance.

A general method for real-time segmentation of moving regions in image sequences entails background subtraction or thresholding the error between an estimation of the image without moving objects and the current image. The various approaches to this problem diverge in the type of background model used and the method used to renew the model. Stauffer, C et al [3] discussed about modeling each pixel as a grouping of Gaussians and uses an on-line approximation to renew the model. The Gaussian distributions of the adaptive mixture model are assessed to determine which are presumably to result from a background process. Each pixel is categorized based on the Gaussian distribution which represents it m ost efficiently is considered part of the background model. This method can be applied to slow lighting changes by slowly adapting the Gaussian values. And can be used with multi-modal distributions caused by shadows, swaying branches, specularities, computer monitors, and other niggling features of the real world.

Wijnhoven et al [5] considered model-based entity detection for traffic surveillance, aiming at object categorization. Within identified regions-of-interest (ROIs) of moving objects in the scene, the orientation of the object is sensed using a gradient direction histogram. For the deliberate orientation, a 3D wire-frame model is applied onto the image data and the finest matching pixel-position is calculated inside the object’s regions-of-interest.

Ismail Haritaoglu et al [19] proposed a low-cost PC based real-time visual surveillance system, called W4, for tracking people and their parts of the body, and supervising their activities in stereo and monochromatic imagery. It operates on grayscale as well as infrared video imagery. This system not uses any color cue. W4 uses a combination of robust tracking techniques, shape analysis, silhouette based body model to point out and track the people and recognize the interaction between people and objects - e.g., putting down objects in the scene, exchanging objects between people, etc. A subsequent system,W4S [20], integrated real-time stereo to overcome the complexities that W4 met with abrupt illumination changes, shadow and occlusion which make tracking much harder in intensity images. An innovative silhouette-based body model is depicted to establish the location of parts of the body when the people are in standard postures. It is a combination of top-down body pose estimation, using a distance transform method that integrates the topology of the human body.
W4S is a visual surveillance system used in real time for identify and track people and monitor their activities in an outdoor environment. It tracks an object by amalgamating real-time stereo calculation with an intensity-based detection and tracking system. W4S does not make use of color cues. But, W4S utilizes a co mbination of shape analysis, stereo and tracking to point out people and their parts (head, feet, hands, and upper body) and produce models of people's appearance with the intention that they can be tracked through interactions such as occlusions. W4S is capable of concurrently tracking several people even with occlusion.

It erects dynamic models of people's movements to respond questions about What are theydoing, and When and Where they act. It builds appearance models of the people it tracks in 2 ½ D so that it can track people (Who?) through occlusion events in the imagery. W4S represents the integration of a real-time stereo (SVM) system with a real-time person detection and tracking system (W4 [19]) to increase its reliability. SVM is a compact, inexpensive real-time device for computing dense stereo range images which was recently developed by SRI. W4S works even in low resolution range maps to continue to track people effectively, because stereo analysis is not considerably affected by sudden illumination changes and shadows, which make tracking much harder in intensity images. Stereo is much helpful in scrutinizing occlusions and other interactions. W4S has the capability to construct a 2 ½ D model of the scene and its human inhabitants by combining a two dimensional cardboard model which signifies the relative positions and body parts size and range as shown in Figure 2.

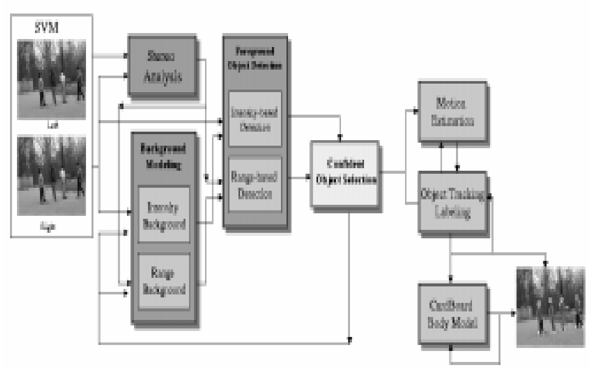

Figure 2: Object detecting system for W4S

Alan J. Lipton et al [26] described an end-to-end method for removing moving targets from a stream of real-time video, sorting them into predefined categories according to image based properties, and after that forcefully tracking them. Moving targets are distinguished using the pixel wise dissimilarity between successive image frames. A classification metric is applied these targets with a chronological consistency constraint to segregate them into three categories: human, background clutter or vehicle. Once segregated, targets are tracked by template matching and temporal differencing combination. The resulting system energetically identifies targets of importance, discards background clutter, and constantly tracks over large distances and periods of time in spite of occlusions, appearance changes and termination of target motion.

To filter out redundant information produced by an array of cameras, and to boost up the response time to forensic events, and to support the human operators with identification of significant events in video a "smart" video surveillance systems is introduced in [14]. The system functions on color as well as gray scale video imagery from a still camera. It can do object detection in outdoor and indoor environments under varying illumination conditions. The classification is done by using the shape of the detected objects. Some advantages of this robust smart video surveillance system are removing shadows, detecting abrupt illumination changes and differentiating left/removed objects.

In [13] a novel approach is presented for multi-object tracking using Minimum Description Length hypothesis. This permits the system to recover from mismatches and provisionally lost tracks. It executes multi-view and multi-category object recognition to detect pedestrians and vehicles in the input images. The 2D object detections are checked for their steadiness with scene geometry and then are renewed to 3D observations.

3. Classification and Tracking of Object

Object tracking is the method of detecting moving objects of interest and plotting its route by analyzing them. Object detection in a video sequence is the method of detecting the moving objects in the frame sequence using digital image processing techniques. Background subtraction is the most commonly used technique for object detection. Background subtraction techniques for object detection from video sequence use the concept of subtracting the background model or a reference model from the current image. The methods considered in tracking of objects use various techniques for building the background model. It has been found that the methods require different time for execution and their performance differs in speed and memory requirements. The techniques involved in these algorithms are based on the intensity values of the pixels constituting the image. The background and illumination changes of the image influence the intensity values to a great extent, ultimately affecting the overall performance. In such situations, these methods fail to give accurate outputs and so there is no single algorithm that performs well in all conditions. An analysis of all these methods based on perturbation detection rate is used to evaluate the performance.

Any tracking method requires an object detection mechanism in each frame or in the first appearance of the object in the video. An ordinary approach for object detection is to use information in a single frame. But, some object detection methods utilize the chronological information computed from a sequence of frames to lessen the number of false detections. This temporal information is usually in the form of frame differencing, which highlights changing regions in consecutive frames. Given the object regions in the image, it is then the tracker's task to perform object correspondence from one frame to the next to generate the tracks. We tabulate several common object detection methods in below Table I [27].

Table 1: Object detection Methods

Categories	Representative Work
Point detectors	Moravec's detector [Moravec 1979], Harris detector [Harris and Stephens 1988], Scale Invariant Feature Transform [Lowe 2004]. Affine Invariant Point Detector [Mikolajczyk and Schmid 2002].
Segmentation	Mean-shift [Comaniciu and Meer 1999], Graph-cut [Shi and Malik 2000], Active contours [Caselles et al. 1995].
Background Modeling	Mixture of Gaussians[Stauffer and Grimson 2000], Eigenbackground[Oliver et al. 2000], Wall flower [Toyama et al. 1999], Dynamic texture background [Monnet et al. 2003].
Supervised Classifiers	Support Vector Machines [Papageorgiou et al. 1998], Neural Networks [Rowley et al. 1998], Adaptive Boosting [Viola et al. 2003].

Rob Wijnhoven et al. [9, 10] presented a patch-based algorithm for the purpose of object classification in video surveillance shown in Figure 3 and 4. Within distinguished regions-of-interest (ROIs) of moving objects in the scene, a feature vector is derived based on template matching of a huge set of image patches. Rather than matching direct image pixels, Gabor-filtered versions of the input image are used at several scales. This approach has been implemented from recent experiments in generic object-recognition tasks. The results have been given for a new typical video surveillance dataset includes over 9,000 object images. In addition, system performance is compared with another existing smaller surveillance dataset. And it is found that with 50 training samples or higher, the detection rate is on the average above 95%. Because of the intrinsic scalability of the algorithm, an embedded system implementation is fine within reach.

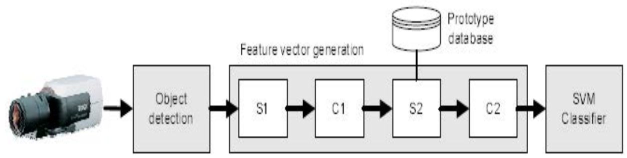

Figure 3: Block diagram for Object classification in camera image

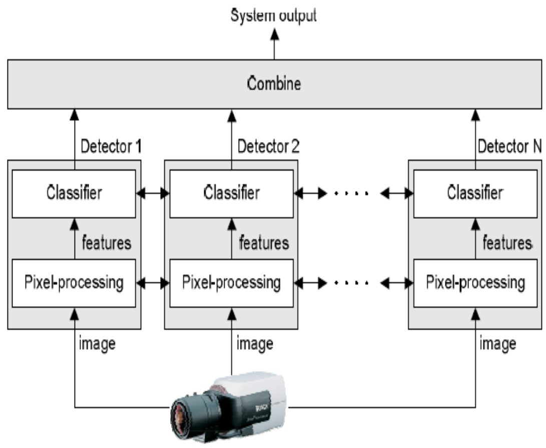

Figure 4: Multiple detectors contained in architecture of Generic object modelling

Several detectors are used in generic object modeling architecture which includes pixel-processing elements and classification systems. They have used a generic architecture as envisaged in Figure above, where detectors can exchange both features extracted at the pixel level and classification results.

Paul Withagen et al [12] proposed a framework where Expectation Maximization (EM) is employed to model the background to detect moving objects. Color histogram of the object is used for Tracking. The probability of the pixel value of the background is calculated using EM method. Advantage of this approach is no thresholds are necessary to refind previously detected objects. This guides to improved object segmentation and simplest detection of occlusion. The main feature of the object is tracked using template matching. Pixels close to the object boundary are detected in every frame with a probability based algorithm and attempt to maintain the color information in the object constant.

In [15], the performance analysis on the background subtraction algorithms were compared the perturbation detection rates of those methods. A custom-made Frame Differencing procedure is proposed. In this work the tracking is based on the mid point of the detected object. The detected object's position is conspired and categorized as human being or vehicles.

In [11], an approach for video surveillance detection of irregular events based on target trajectory analysis is presented. The methodology pursues a sequence: Detection, Tracking and Identification. The detection step is done by using the color constancy principle using an adaptive background subtraction technique with shadow elimination model. The tracking uses a direct and inverse matrix matching process. In the identification stage local motion properties are expressed by elliptic Fourier descriptors.

4. Automatic Video Surveillance

Real-time segmentation of moving regions is an elemental step in several vision systems including human-machine interface, automated visual surveillance and very low-bandwidth telecommunications. A typical method used here is background subtraction. Numerous background models have been brought in to handle different problems. Pixel based Multi-colour background model proposed by Grimson et al [3, 30, and 31] is one of the successful solutions to these problems. However, this method suffers from slow learning at the beginning, especially in busy environments and it couldn't differentiate between moving objects and moving shadows. P. KaewTraKulPong et al [29] presented a method which improves this adaptive background mixture model. By reinvestigating the update equations, we utilize different equations at different phases. This allows our system learns faster and more accurately as well as adapt effectively to changing environments. A shadow detection scheme is established. It is derived from a computational colour space that makes use of the background model specified here.

Karan Gupta et al [16] presented a work based on the concepts of dynamic template matching and frame differencing to employ a robust automated object tracking system. In their implementation an indistinct industrial camera has been used to seize the video frames and pursue an object. Frame differencing is used frame-by-frame basis in a moving object to detect with high accuracy and efficiency. After the object has been detected it is tracked by utilizing an efficient Template Matching algorithm. The patterns used for the matching purposes are generated with dynamism. This guarantees that any change in the pose of the object does not obstruct the tracking procedure. To mechanize the tracking process the camera is mounted on a pan-tilt arrangement, which is harmonized with a tracking algorithm. Whenever the object being tracked shift out of the viewing range of the camera, the pan-tilt setup is robotically adjusted to move the camera to keep the object in view. The system is competent of handling object's entry and exit. This tracking system is cost effectual and used as an automated video conferencing system and video surveillance tool.

Three key steps are there in implementation of this tracking system:

1) Detection of interesting moving objects by Frame Differencing
2) Tracking of such objects from frame to frame by Dynamic Template Matching
3) Analysis of object tracks to automate the pan-tilt mechanism

Yuhua Zheng et al [25] have presented an automatic object detecting and tracking algorithm by using particle swarm optimization based method. PSO is a searching algorithm stimulated by the behaviors of social insect in the nature. To detect objects, a flow of boosted classifiers based on Haar-like features is trained and used. To improve the searching effectiveness, initially the object model is projected into a high-dimensional feature space, and then PSO-based algorithm is used to search over the high-dimensional space and congregate to some global features of the object. After that, a Bayesian filter is used to recognize the best match with the peak possibility amid these candidates under the restraint of object motion estimation. This PSO based algorithm considers even the object motion estimation to speed up the searching procedure.

Axel Baumann et al [17] provided a systematic review of measures and evaluates their effectiveness for specific features like segmentation, event detection and tracking. This review focuses on normalization issues, representativeness and robustness. A software framework is established for continuous evaluation and documentation of the performance of video surveillance systems. A new set of representative measures is projected as a primary part of an evaluation framework.

F. Dufaux et al [23, 24] proposed a region-based transform-domain scrambling technique. As a first step, video data defer regions of interest (ROI), such as face or license plates that are about to contain private-sensitive information. These are then twisted to obscure content. The approach is a standard one and can be used to any transform-coding method, such as might be based on discrete wavelet transform (DWT) or discrete cosine transform (DCT). The obscuring content is done in the transform domain by pseudo-randomly flipping the sign of transform coefficients during encoding. This method facilitates the level of distortion to be attuned, from mild uncertainty to complete noise. Consequently, the scene can be implicit even if individuals in it cannot be located out. The obscured content depends upon a private encryption key and it is reversible. The key can be entrusted to law-enforcement authorities to authorize unlocking and viewing the scene in clear.

5. Multimedia Surveillance (MSS)

MSS utilize assorted number of related media streams, each of which has a different assurance level to attain numerous surveillance tasks. For example, the system designer may have a higher poise in the video stream balanced to the audio stream for detecting humans running events. The assurance level of streams is usually precalculated based on their earlier accuracy. This traditional approach is difficult especially when we insert a new stream in the system with no knowledge of its prior history. Pradeep et al [18] proposed a novel method which vigorously computes the confidence level of new streams based on the fact whether it provides evidence which concurs or contradicts with the already trusted streams.

Atrey et al [21] **p**roposed a framework for "when" and "how" to absorb the information obtained from multiple sources to facilitate detecting events in multimedia surveillance systems. And it addresses about determining the finest subset of sensor (streams). The proposed method espouses a hierarchical probabilistic assimilation approach and carries out assimilation of information at three distinct levels - media stream level, atomic event level and compound event level. To detect an event, this framework uses the media streams available at the current instant and utilizes two important properties of them namely, accrued past history of whether they have been providing concurring or contradictory evidences, and the system designer's poise in them.

In [22] a framework is proposed which uses a dynamic programming based method that finds the finest subset of media streams derived from three different criteria; first, by maximizing the probability of the occurrence of event with a specified minimum confidence and a specified maximum cost; second, by maximizing the confidence in the subset with a specified minimum probability of the occurrence of event and a specified maximum cost; and third, by minimizing the cost of employing the subset with a specified minimum probability of the occurrence of event and a specified minimum confidence. The proposed dynamic programming based method allows for a tradeoff among the above-mentioned three criteria, and put forwards the flexibility to evaluate whether any one set of media streams which costs less would be better than any other set of media streams which costs high, or any one set of media streams with high confidence would be better than any other set of media streams with less confidence.

Rama et al [28] addressed the problem of how to choose the most favorable number of sensors and how to decide their location in a given monitored area for MSS. For that they have proposed a n ew performance metric for achieving the given surveillance task using varied sensors and presented a n ew design methodology based on that metric which can help attain the optimal combination of sensors and additionally their placement in a given surveyed area. The same measure can be used to analyze the dilapidation in system's performance regarding the failure of various sensors. They have built a s urveillance system using the finest set of sensors attained based on the proposed design methodology.

With the escalating use of audio sensors in surveillance and monitoring applications, event detection using audio streams has materialized as a key research problem. Atrey et al [21] presented a top-down event detection approach for audio based event detection for surveillance. The proposed approach initially sorts a given audio frame into non-vocal and vocal events, and next carry outs advance classification into excited and normal events. Gaussian Mixture Model is used to optimize the parameters for four different audio features LPC (Linear Predictive Cepstral Coefficients), ZCR (Zero Crossing Rate), LPCC (Linear Predictive Cepstral Coefficients) and LFCC (Log Frequency Cepstral Coefficients). The results show that the proposed hierarchical event detection approach works significantly better than the single level approach.

The system consists of two stages - offline training (or event modeling) and online testing (event detection) as shown in Figure 5.

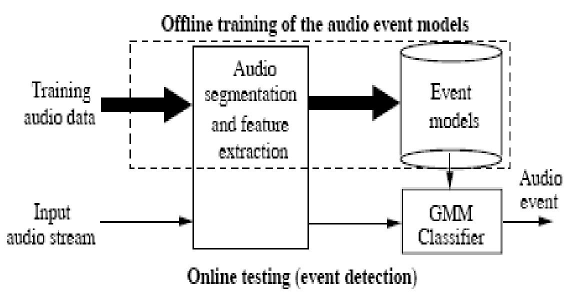

Figure 5: Event detection system using Audio Frame

As shown in figure 6, at the top level (Level 0), each input audio frame is classified as the foreground or the background. The background is the environment noise which represents ‘no event’ and is ignored. The foreground that represents the events, are further categorized into two classes - vocal and nonvocal (level 1). At the next level (2), both vocal and nonvocal events are further classified into normal and the excited events. Finally, at the last level (3), the footsteps sequences are classified as walking or running based on the rate of their occurrence in a specified time interval.

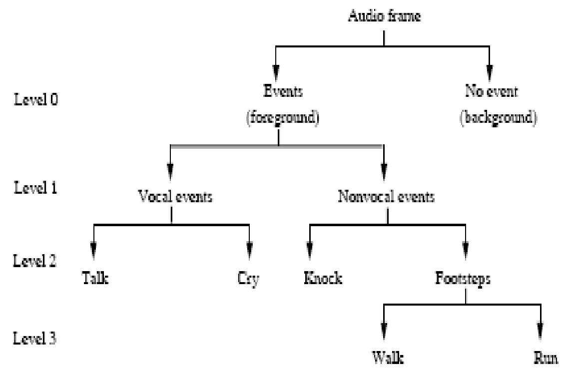

Figure 6: A top-down approach for event detection using Audio.

6. Conclusions

This paper reviews and exploits the existing developments and different types of video surveillance systems which are used for object tracking, behavior analysis, motion analysis and behavior understanding. The inspiration of writing a survey paper on this topic is to evaluate and reach insight in visual surveillance systems from a big picture first. This enables us to understand and answer the questions like: What are the Developments and different stages involved in a visual surveillance system; how to detect and dissect behavior and intent; etc.

6. References

[1] Ching-Kai Huang and Tsuhan Chen, “Motion Activated Video Surveillance Using TI DSP”, DSPS FEST '99, Houston, Texas, August 4-6, 1999.

[2] Lun Zhang Li, S.Z. Xiaotong Yuan Shiming Xiang, Beijing, “Real-time Object Classification in Video Surveillance Based on Appearance Learning“, Computer Vision and Pattern Recognition, 2007, IEEE Conference Issue Date: 17-22 June 2007, Print ISBN: 1-4244-1180-7, pp.1-8.

[3] Stauffer. C, Grimson W.E.L, “Adaptive background mixture models for real-time tracking”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Volume 2., Ft. 1999, pp. 252–258.

[4] K. Srinivasan K. Porkumaran, G. Sainarayanan, “Intelligent human body tracking, modeling, and activity analysis of video surveillance system: A survey”, Journal of Convergence in Engineering, Technology and Science Vol.1 December 2009, pp. 1-8.

[5] Wijnhoven, R., de With, P, "3D wire-frame object-modeling experiments for video surveillance", Proc. of 27th Symposium on Information Theory in the Benelux, 2006, pp. 101–108.
[6] Massimo Piccardi, "Background subtraction techniques: a review", IEEE International Conference on Systems, man and cybernets 2004.
[7] Muller-Schneiders, S., Jager, T., Loos, H., Niem, W., "Performance evaluation of a real time video surveillance system", Proc. of 2nd Joint IEEE Int. Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS), 2005, pp. 137–144.
[8] Nan Lu et. al, "An Improved Motion Detection Method for Real-Time Surveillance", Proc. of IAENG International Journal of Computer Science, Issue 1, 2008, pp.1-16.
[9] R. G. J. Wijnhoven, P. H. N. de With, "Experiments With Patch-Based Object Classification", IEEE Int. Conf. on Advanced Video and Signal based Surveillance, 2007.
[10] R. Wijnhoven, Peter H.N. de With, "Patch-Based Experiments with Object Classification in Video Surveillance", ACIVS 2007, LNCS 4678, 2007, pp. 285–296.
[11] Mohamed Dahmane, Jean Meunier, "Demo: Video-Surveillance application with Self-Organizing Maps".
[12] Paul Withagen, Klamer Schutte, Frans Groen, "Object Detection and Tracking Using a Likelihood Based Approach", Proceedings of the ASCI 2002 conference, Lochem, The Netherlands, June 19-21 2002, pp. 248-253.
[13] Bastian Leibe, Konrad Schindler, Nico Cornelis, Luc Van Gool, "Coupled Object Detection and Tracking from Static Cameras and Moving Vehicles".
[14] Yigithan Dedeoglu, "Moving Object Detection, Tracking And Classification For Smart Video Surveillance" 2004.
[15] Rajitha T V, "Object Tracking From Video Sequence", June 2008.
[16] Karan Gupta, Anjali V. Kulkarni, "Implementation of an Automated Single Camera Object Tracking System Using Frame Differencing and Dynamic Template Matching".
[17] Axel Baumann, Marco Boltz, Julia Ebling, Matthias Koenig, Hartmut S. Loos, Marcel Merkel, Wolfgang Niem, Jan KarlWarzelhan, Jie Yu "A Review and Comparison of Measures for Automatic Video Surveillance Systems", Hindawi Publishing Corporation EURASIP Journal on Image and Video Processing Volume 2008, Article ID 824726, 30 pages doi:10.1155/2008/824726.
[18] P. K. Atrey, Mohan S. Kankanhalli, Abdulmotaleb El Saddik, "Confidence Building Among Correlated Streams in Multimedia Surveillance Systems", MMM 2007, LNCS 4352, Part II, 2007, pp. 155–164.
[19] Ismail Haritaoglu, Harwood, D., Davis, L, "W4: real-time surveillance of people and their activities", IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 22, 2000, pp. 809–830.
[20] Ismail Haritaoglu, David Harwood, Larry S. Davis, "W4S: A Real-Time System for Detecting and Tracking People in 2 ½ D".
[21] Atrey, P.K., Maddage, N.C., Kankanhalli, M.S., "Audio based event detection for multimedia surveillance", IEEE International Conference on Acoustics, Speech and Signal Processing, 2006, pp 813–816.
[22] Atrey, P.K., Kankanhalli, M.S., Jain, R., "A framework for information assimilation in multimedia surveillance systems", ACM Multimedia Systems Journal, 2006.
[23] F. Dufaux, M. Ouaret, Y. Abdeljaoued, A. Navarro, F. Vergnenegre, T. Ebrahimi, "Privacy Enabling Technology for Video Surveillance", Proc. SPIE 6250, 2006.
[24] F. Dufaux, T. Ebrahimi, "Scrambling for Video Surveillance with Privacy", Proc. IEEE Workshop on Privacy Research In Vision, New York, NY, June 2006.
[25] Yuhua Zheng and Yan Meng, "Adaptive Object Tracking using Particle Swarm Optimization", Proceedings of the 2007 IEEE International Symposium on Computational Intelligence in Robotics and Automation, Jacksonville, FL, USA, June 20-23, 2007, pp 43-49.
[26] Alan J. Lipton, Hironobu Fujiyoshi, Raju S. Patil, "Moving Target Classification and Tracking from Real-time Video".
[27] Yilmaz, A., Javed, O., and Shah, M. 2006. "Object Tracking: A Survey", ACM Comput. Surv. 38, 4, Article 13 Dec 2006.
[28] Rama, K.G.S., Atrey, P.K., Singh, V.K., Ramakrishnan, K., Kankanhalli, M.S., "A Design Methodology For Selection and Placement of Sensors in Multimedia Surveillance Systems", The 4th ACM International Workshop on Video Surveillance and Sensor Networks, Santa Barbara, CA, USA 2006.
[29] P. KaewTraKulPong, R. Bowden, "An Improved Adaptive Background Mixture Model for Real-time Tracking with Shadow Detection", Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems, AVBS01, Sept 2001.
[30] Grimson Wel, Stauffer C. Romano R. Lee L., "Using Adaptive Tracking To Classify And Monitor Activities In A Site", Proceedings.1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1998.
[31] Stauffer C, Grimson W. E. L., "Learning patterns of activity using real-time tracking", IEEE Transactions on Pattern Analysis & Machine Intelligence, 2000. Vol. 22 No. 8, p. 747-57.

C. Lakshmi Devasena is presently doing PhD in Karpagam University, Coimbatore, Tamilnadu, India. She has completed MPhil (computer science) in the year 2008 and MCA degree in 1997 and B.Sc (Mathematics) in 1994. She has seven and half years of teaching experience and two years of industrial experience. Area of research area is Image processing and Data mining. At Present, she is working as Lecturer in Karpagam University, Coimbatore.

M. Hemalatha completed MCA MPhil., PhD in Computer Science and currently working as an Assistant Professor and Head, Department of software systems in Karpagam University. Ten years of Experience in teaching and published Twenty seven papers in International Journals and also presented seventy papers in various National conferences and one international conferences Area of research is Data mining, Software Engineering, bioinformatics, Neural Network. She is also reviewer in several National and International journals.