

SMART SURVEILLANCE APPLICATIONS, TECHNOLOGIES & IMPLICATION

Dinesh Pandu Ranga Gupta Chakka.
Computer Science And Engineering
Lovely professional University Jalandhar, India.
chakkaguptassdj@gmail.com

Venkata Sai Pradeep Sunku.
Computer Science And Engineering Lovely professional University Jalandhar, India.
Saipradeep006@gmail.com

Krishna Chaitanya Atchutuni
Computer Science And Engineering Lovely professional University Jalandhar, India.
Chaitu.achutini@gmail.com

Swati
Computer Science And Engineering Lovely professional University Jalandhar, India.
rampalswati@gmail.com

ABSTRACT

Smart surveillance, is the use of automatic video analysis technologies in video surveillance application. This paper attempts to answer a number of question about smart surveillance: what are the application of smart surveillance? What are the system architecture for smart surveillance? What are the key technologies? What are the some of the key technical challenges? And what are the implication of smart surveillance, both to security and privacy?

INTRODUCTION

Recent world events have created a shift in the security paradigm from “investigation of incidents” to “prevention of potentially catastrophic incidents”. Existing digital video surveillance system provide the infrastructure only to capture, store and distribute video, while leaving the task of threat detection exclusively to human operators. Human monitoring of surveillance video is a very labour-intensive task. It is generally agreed that watching video feeds requires a higher level of visual attention than most every day tasks. It is generally agreed that watching video feeds requires a higher level of visual attention than most every day tasks Specifically vigilance, the ability to hold attention and to react to rarely occurring events, is extremely demanding and prone to error due to lapses in attention and to react to rarely occurring events, is extremely demanding and prone to error due to lapses in attention[12]. One of the conclusion of recent study by the us national institute of justice[6], into the effectiveness of human monitoring of surveillance video, is as follows

“These studies demonstrated that such a task [... manually detecting events in surveillance video] even when assigned to a person who is dedicated and well-intentioned, will not support an effective security system. After only 20 minutes of watching and evaluating monitor screens, the attention of most individuals has degenerated to well below acceptable level. Monitor screens, the attention of most individuals has degenerated to well below acceptable level. Monitoring video screens is both boring and mesmerizing. There are no intellectually engaging stimuli, such as when watching a television program ”.

Clearly today’s video surveillance system while providing the basic functionality fall short of providing the level of information need to change the security paradigm from “investigation to pre-emption”. Automatic visual analysis

technologies can move today's video surveillance system from the investigative to preventive paradigm. Smart surveillance system provide a number of advantages over traditional video surveillance system, including

“The ability to pre-empt incidents - - through real time alarms for suspicious behaviors”.

“Enhanced forensic capabilities through based video retrieval”

Situation awareness through joint awareness of location, identify and activity of objects in the monitored space”.

Section 2 Provides a short introduction to various application of smart surveillance system. Section 3 Discusses the architectures for smart surveillance system. Section 4 Presents the key technologies for smart surveillance system. Section 5 Briefly discusses the challenges in smart surveillance. Section 6 Discusses the implication of smart surveillance technologies. Conclusions are presented in section 7

1. Application of smart surveillance

In this section we describe a few application of smart surveillance technology. In this section, we describe a few applications. We group the application into three broad categories, real time alerts, automatic forensic video retrieval, and situation awareness.

2.1 Real time alerts: there are two types of alerts in smart surveillance system user defined alerts Automatic unusual activity alerts.

2.1.1 User defined alerts: Here the system is required to recognize a variety of user defined events that occur in the monitored space and notify the user in real time, thus providing the user with an opportunity to evaluate the situation and take preventive action if necessary. Following are some typical events.

1. Generic alerts:

These are alerts which depends solely on the movement properties of objects within the monitored space following are the few examples.

1. Motion detection (it will detects movement of any object in specified zone)
2. Motion characteristic detection: (it will detects the motion of the object and it include the specific direction of object movement)
3. Abandoned object alert:(This detects objects which are abandoned)
4. Object removal(this will detects movements of a user-specified object that is not expected to move).

2.Class specific alerts:

These are alerts which use the type of object in additional to the object movement properties

Example: type specific movement detection consider a camera that is monitoring runways at an airport. In such a science, the system could provide an alert on the presence or movement of people on the tarmac but not those of aircrafts

Statistics: example applications include, alerts based on people counts(e.g. more than one person in security locker) or people densities (e.g. discotheque crowded beyond an acceptable level).

3.Behavioral alerts:

These alerts are generated based on adherence to, or deviation from, learnt models of motion patterns. Such models are typically trained by analyzing movement patterns over extended periods of time. These alerts tend to be very application specific and use a signification amount of context information for example : detecting shopping groups at retail checkout counters, and alerting the store manager when the length of the queue at a counter exceeds a specified number

Detecting suspicious behavior in parking lots for example a person stopping and trying to open multiple cars

4.High value video capture:

This is an application which augment real time alerts by capturing selected clips of video based on prespecified criteria this becomes highly relevant in the context of smart camera network which use wireless communication

2.1.2 Automatic unusual activity alerts:

Unlike the user defined alerts here the system generates alerts when it detects activity that deviates from the norm. the smart surveillance system achive that is monitoring a street learns that vehicles move about on the road and people move about on the side walk'. Based on this pattern the system will provide an alert when a car drive on the side walk. Such unusual activity detection Is the key t effective smart surveillance as all the events of interest con not be manually specified by the user.

2.2 Automatic forensic video retrieval(AFVR):

The capability to support forensic video retrieval is based on the rich video index generated by automatic tracking technology. This is a critical value-add from using smart surveillance technologies. Typically the index consists of such measurement as object sharp, size and appearance information, temporal trajectories of objects over time object type information, in some cases specific object identification information . in advanced system, the index may contain abject activity information. The washington, DC sniper incident in a prime example of where AFVR could be a break-through technology. During the incident the investigative agencies had access to hundreds of hours of video surveillance footage draw from a wide variety of surveillance cameras covering the areas in the vicinity of various incidents. However, the task of manually sifting through hundreds of hours of video for investigative purposes is almost impossible. However if the collection of videos were indexed using visual analysis , it would enable the following ways of retrieving the video.

1. SPATION-TEMPRAL VIEDO RETRIVAL:

An example query in this class would be, "retrieve all clops of video where a "blue car" drove in front of the "7/11 store on 23rd street" between the 26thof July 2pm and 27th of July 9am" at speed>25mph".

2. SURVEILLANCE VIEDO MINING:

In case of the washington sniper incident, the surveillance video mining application would attempt to present the users with a set of potential movement patterns I of cars over a set of camera covering multiple incident location, this would enable the investigative agencies to answer question like "was there a single car that appeared in all of the incident location?"

2.3 SITUATION AWARENESS:

Ensuring total security at a facility requires system that can perpetually track the identity, location and activity of people and vehicle within the monitored space. For example ,the existing surveillance technology cannot answer question such as did a single person loiter around near a high security building on multiple occasions? Such perpetual tracking can be the basis for very levels of security typically surveillance system have focused on tracking location and activity, while biometrics system have focused on identifying individuals. As smart surveillance technologies mature it becomes possible to address all these three key challenges in a single unified frame work giving rise to, joint location identity and activity awareness, which when combined with the application context becomes the basis for situation awareness.

3. Smart surveillance architectures:

In this section how smart surveillance technologies are incorporated into a complete surveillance system we discussed three different types of smart surveillance architectures

3.1 Basic smart surveillance architecture(BSSA):

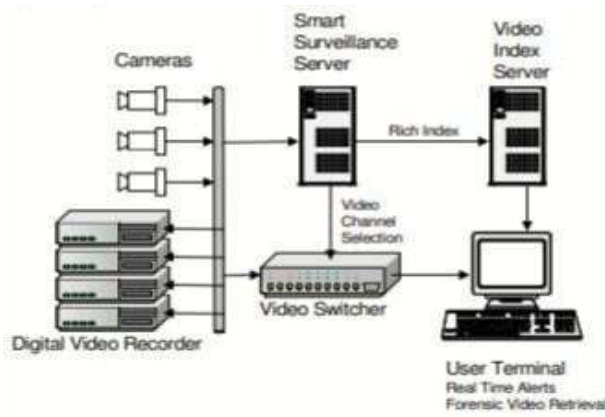


Figure 1 Block Diagram of smart surveillance system.

Figure 1 show the block diagram of smart surveillance system. The outputs of video cameras are recorded digitally and simultaneously analyzed by the smart surveillance server, which produces real time alerts and a rich video index. The types and parameters of the alerts are user configurable.

The user can use the rich index to retrieve video from the archive for forensic. In the BSSA 3 the video recording and analysis is centralized requiring the cameras to be wired to a central location. The role of automatic visual in the BSSA is primarily analytical.

3.2 ACTIVE SMART SURVEILLANCE ARCHITECTURE(ASSA):

Figure 2 show the block diagram for an active smart surveillance architecture. The key difference between this as BSSA is the additional of active camera controls. Here the automatic visual analysis is used not only to understand what is going on in the scene but also “to selectively pay more attention” to automatically detected activities or event of interest. The ASSA could be used for many different application. Below we describe two example

1. Face cataloger : this is a system which aims to non intrusively acquire a high- resolution face images of all people passing through a space, here ASSA detects and track people and uses the active cameras to zoom in and acquire high resolution face pictures.
2. Multi-scale video: This is a system which automatically allocates higher resolution to portions of the scene which have certain predetermined types of activity for example, all cars that are moving with high speed through a parking lot may be imaged at a higher resolution through the active cameras.

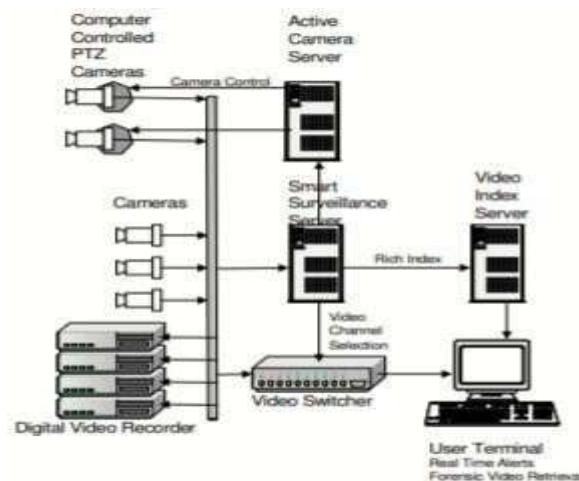


Figure 2 Block diagram of an Active Smart.

The key technical difference between the ASSSA and BSSA are

Active camera resource management : this include techniques for deciding which of the active cameras are to be used to meet the overall goals of the system

Active camera image analysis: this include analysis of the camera images in order to control the movement and zoom of the cameras.

DISTRIBUTED SMART SURVEILLANCE ARCHITECTURE(DSSA)

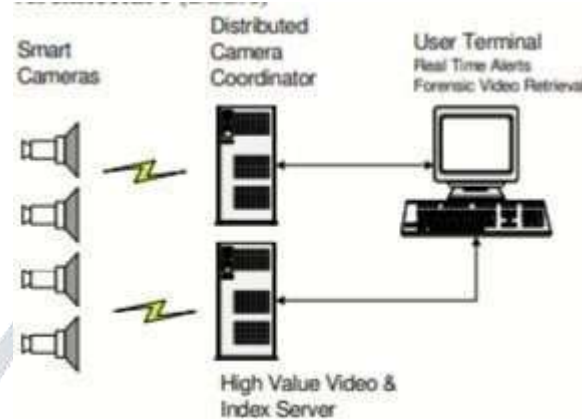


Figure 3 Block diagram of a distributed smart surveillance architecture using smart cameras

One of the biggest cost in deploying a surveillance system is the infrastructure(e.g., wiring) required to move the video from the camera to a central location where it can be analyzed and stored. Further, the wiring capacity is not easily scalable and because of involvement of manual labour, instantaneous/ portable installation of network is difficult.

There is an increasing trends to building camera which in addition to generating images also analyzed then with on board processing. Such camera are typically called smart camera. The goal of a smart camera system is to minimize the cost of deployment. In such architecture(fig3), the camera would typically use wireless communication to coordinate with a central camera coordinator. A smart camera would use the automatic visual analysis to determine how to use the limited storage and bandwidth to effectively transmit only “high value video” to the server.

4.TECHNOLOGIES FOR SMART SURVEILLANCE

Among the three architectures presented presented in the previous section, the BSSA is likely to be the most ubiquitous in the near future. The BSSA is an enhancement of the architecture of current day surveillance system, which have cameras mounted in different parts of a facility, all of which are wired to a central room. In this section, we will discuss various technologies that enable the BSSA and discuss the challenges involved as such system get widely deployed.

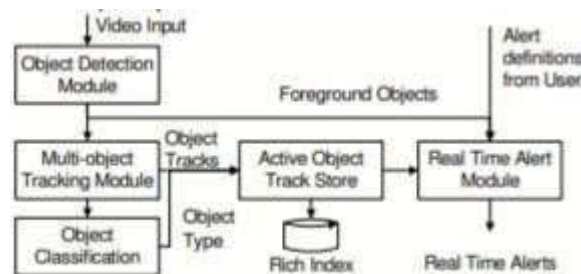


Figure 4 Internal structure of the smart surveillance engine.

4.1 Object detection:

Most of the work on object detection relies heavily on the assumption of a static camera[9]. There is some work which has looked at detecting independent motion in moving camera images where the camera motion is well modelled[18]. Our approach to object detection has been two pronged each of which we discuss briefly

4.1.1 Adaptive background subtraction with healing: The background subtraction module combine evidence from difference in colour, texture, and motion. The use of multiple modalities improve the detection of objects in cluttered environments. The resulting saliency map is smoothed using morphological operators and then small holes and blobs are mechanisms to handle changing ambient conditions and scene composition. First, it continually updates its overall RGB channel noise parameters to compensate for changing light levels. Second, it estimates and corrects. Thirdly it maintains a map of high activity regions and slowly updates its background model only in areas deemed as relatively quiescent. Finally, it automatically eliminates occasional spurious foreground object based on their motion patterns.

4.1.2 Salient motion detection: This is a complementary approach to background subtraction. Here we approach the problem from a motion filtering perspective. Consider the following figure 5. on the left show a scene where a person is walking in from of a bush which is waving in the wind. The next figure 5 show the output of a traditional background subtraction algorithm (which per its design correctly classifies the entire bush as a moving object) . however , in this situation, we are interested in detecting the person as opposed to the moving bush. Our approach uses optical flow as the basis for detecting salient motion. We use a temporal window of n frames (typically 10-15) to assess the coherence of optical flow at each pixel over the entire temporal window. Pixels with coherent optical flow are labelled as candidates. The candidates from the motion filtering are then subjected to a region growing process to obtain the final detection.

4.2 Multi-object tracking:

Multi-object tracking attempts to associate object with one another over time, by using a combination of the object appearance and movement characteristics. This has been a very active area of research [2,3,4,5,8] in the past several years. Our approach to multi- object tracking has focused heavily on handling occlusions. Following is a brief description of our approach. The multi-object blob tracking relies on appearance models which are image-based templates of object appearance. New appearance models are created when an object enters a scene. In every new frame, each of the existing tracks is used to try to explain the foreground pixels.

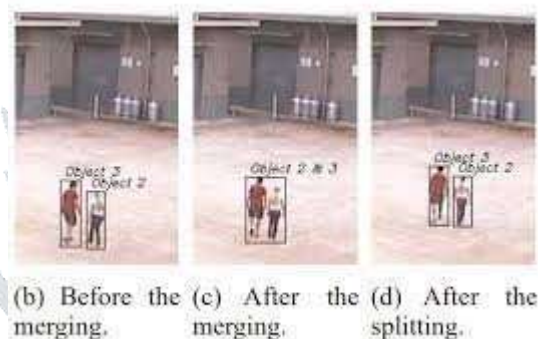


Figure 5 Occlusion handling in object tracking: Left Original image with occlusion. Middle: Object being tracked by the system. Right: Classification of pixels into models.

4.3 Object classification:

Moving fore ground object are classified into relevant categories. Statistics about the appearance, shape, and motion of moving objects can be used to quickly distinguish people, vehicles, carts, animals, doors opening/closing , tress moving in the breeze, etc., our system classifies object into vehicles, individuals, and groups of people based on shape features, recurrent motion measurement, speed and direction of motion. From a small set of training

4.4 Real time alert module:

The real time alert module uses the information produced by the other modules, namely, object detection, tracking and classification to detect user specified alerts. The key feature of the real time alert module is its extensible design it show the generic structured of the real time alert module. This structure is instantiated by most of the alert as an example . the following processes are instantiated when the user specifies a directional motion alert.

Directional motion alert manager: Each motion alert user definition instantiates a directional motion alert manager, which is responsible for ensuring correct monitoring of the scene. The alert manager ensures that for every object being tracked there is a corresponding object track observer that is instantiated. And that the object track observer is deleted immediately upon the exit of object from the scene

Direction motion object track observer: This is the process that is charged with the job of measuring the direction of motion of the object and comparing it to the user specified direction. When ever the object motion direction matches the user specified direction, the object track observer issues a real time alert. The application uses the alert to signal the user that one of the specified alert conditions has been met.

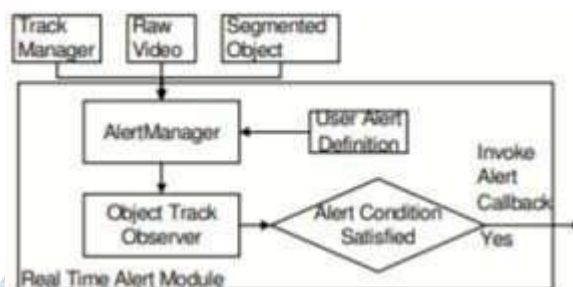


Figure 7 internal structure of the real time alert module.

The extra nature of the object track observer depends on the particular alert it is implementing however the general structure of many types of alerts is similar to that describe above.

5. challenges:

1.technical challenges: these are the a number of technical challenges that still need to be addressed in the underlying visual analysis technologies. These include challenges in robust object detection, tracing object in crowded environments, challenges in tracking articulated bodies for activity understanding, combining biometric technologies like face recognition with surveillance to achieve situation awareness

2.challenges in performance evaluation: this is a very signification challenge in smart surveillance system evaluating annotation is very expensive ad tedious process. Additionally, there can be significant errors in annotation. All of these issues make performance evaluation a significant challenge

6. conclusions:

While it is difficult to foresee a future where the surveillance of the monitored space is completely automatic, there is clearly an urgent need to augment the exiting surveillance technology with better tools to aid efficacy of the human operator. With the increasing availability of the inexpensive computing , video infrastructure and better. Video analysis technologies smart surveillance system will completely replace existing surveillance system. The degree of smartness” will vary with the level of security offered by such system

7. Acknowledgments:

All the technical discussions in this paper are based on ongoing work in visual tracking at IBM research. The authors wish to acknowledge the people vision project at IBM T.J.WASTSON research centre.

REFERENCES

1. Lisa Brown and Yingli Tian, "Comparative Study of Coarse Head Pose Estimation", *IEEE Workshop on Motion and Video Computing*, Dec. 5-6, 2002[1].
2. Anjum Ali and J. K. Aggarwal, "Segmentation and Recognition of Continuous Human Activity", *IEEE Workshop on Detection and Recognition of Events in Video*, 2001[2].
3. Collins, Lipton, Fujiyoshi and Kanade, "Algorithms for cooperative multisensor surveillance", *Proc. IEEE*, vol. 89, no. 10, Oct. 2001[3].

4. D. Comaniciu, V. Ramesh and P. Meer, "Real-Time Tracking of Non-Rigid Objects using Mean Shift", *IEEE Conf. Computer Vision and Pattern Recognition (CVPR'00)*, vol. 2, pp. 142-149, 2000
5. Trevor Darrell, David Demirdjian, Neal Checka and Pedro Felzenszwalb, "Plan-View Trajectory Estimation with Dense Stereo Background Models", *ICCV 2001*, pp. 628-635
6. Mary W. Green, "The Appropriate and Effective Use of Security Technologies in U.S. Schools", A *Guide for Schools and Law Enforcement Agencies Sandia National Laboratories*, September 1999
7. A. Hampapur et al., "Face Cataloger: Multi-Scale Imaging for Relating Identity to Location", *IEEE International Conference on Advanced Video and Signal Based Surveillance*, July 03.
8. Haritaoglu and Flickner, "Detection and Tracking of Shopping Groups in Stores", *CVPR*, 2001.
9. T. Horprasert, D. Harwood and L. Davis, "A Statistical Approach for Real-Time Robust Background Subtraction and Shadow Detection", *Proceedings of IEEE Frame-Rate Workshop*, 1999.
10. Biometrics: Personal Identification in Networked Society, Kluwer Academic, 1999, ISBN 0792383451.

