Abstract—Recently, Crowd analysis became the hotspot field in research. Automated analysis of crowd activities using surveillance video is an important means for security. In public venues the crowd size is the key indicator for the safety and stability. Now a day many researchers show there interest in this area as there is a growing need of detecting the abnormalities of crowd by using various methods such as density, behavior, tracking and motion estimation. In this paper we present the four main components of the crowd analysis that are density estimation, behavior recognition, crowd tracking and motion estimation and also the main key for security is counting of people.

Keywords—Crowd analysis, Components of analysis, Counting of People, Video Surveillance

INTRODUCTION

Video Surveillance has become a hotspot area with the upcoming technologies and the increasing need for security. The video surveillance application use the crowd analysis for automatic detection of anomalies and alarm. The main task of the video based detection scheme consists of recognizing and tracking of human. The venues such as airports, railway station, sports and rallies specially require monitoring.

Video surveillance systems are widespread and common in many environments. More recently, governments’ agencies, businesses, and even schools are turning toward video surveillance as a means to increase public security. The important element of video surveillance is not only on the placement of the cameras for human eyes, but also for fully automated surveillance activities[1]. Applications of video surveillance include car and pedestrian traffic monitoring, human activity surveillance for unusual activity detection, people counting, etc. A typical surveillance application consists of three buildings blocks: moving object detection, object tracking and higher level motion analysis. Multimedia systems can provide surveillance coverage across a wide area, ensuring object visibility over a large range. Several video surveillance products are available in the market for office and home security as well as remote surveillance. They monitor a home, an office, or any location of interest, capturing motion events using webcams or camcorders and detect abnormalities [2]. In the case of webcams, the visual data is saved into compressed or uncompressed video clips, and the system trigger various alerts such as sending an e-mail.

The remaining paper is organized as follows : Section I Video surveillance system, Section II Components of Crowd analysis. Section III Counting people technique, IV Some benefits and drawbacks of video surveillance.

I. VIDEO SURVEILLANCE SYSTEM

The Surveillance system consist of :
• Detection
• Tracking
• Classification
• Person Recognition

Detection : The main difficulty of detection is that, the background undergoes a continual change (example: clouds passing by, branches of trees moving with the wind). Detecting objects is a critical capability for surveillance. From the perspective of a human intelligence analysis, the most critical challenge in video based surveillance is interpreting the data to detect events of interest and identify trends.

There are two approaches to object detection:
1. background subtraction
2. salient motion detection.

Tracking : The purpose of tracking is to determine the information of each target present in the scene. It is also challenging task in surveillance.

Classification : For the classification task three main questions must be answered, namely: which classes should be considered, which features best separate these classes and which classifiers best adapt to the previous choices?

Person Recognition : The recognition process aimed to recognizing the objects in a short term period, i.e. targets that become close for a few seconds or targets that merge for a few seconds and then split again[2].
II. COMPONENTS OF CROWD ANALYSIS

Crowd analysis is the hotspot area of research. Crowd can be a group of individuals in a general term, the behavior of the crowd has a collective characteristic such as ‘angry crowd’ and ‘peaceful crowd’. The components of crowd analysis consist of crowd density estimation, crowd motion detection, crowd tracking, and crowd behavior understanding.

A. Crowd density estimation

Crowd density estimation is the challenging task for visual surveillance. Due to presence of high risk of degeneration, the safety of public events has large crowds and it has always been major concern. Video analysis techniques are becoming increasingly popular in the visual surveillance of public areas because of their great efficiency in gathering information and low cost in human resource. With the rapid economic development and increasing people's social activities, the flow density of large malls, supermarkets and places such as subway station is growing more and more serious and brings security risks by the crowd congestion. Two main target of crowd density estimation:

1. Providing an approximation number of how many people are in target scene.
2. Providing a range of people in the crowd i.e. determine the density in broad classes.

There are two ways to estimate the crowd density. By using the Haar features i.e. the digital image features used in object recognition and by the texture feature extraction.

1. Pre-processing

In Pre-processing video frames are taken as an input and these frames are pre-processed. Pre-processing is step to convert the color image to gray image as time required for processing the gray image is less than the color image.

2. Background Generation

This method is used to distinguish foreground objects from the stationary background. To achieve this first N-frames are combined to generate the strong background. This background is then subtracted from the input frame to extract foreground of the input.

3. People Detection

Human blobs are detected using AdaBoost detector and Periodic Dynamicity Estimation is performed to eliminate detected non-human movements if any. Then the crowd density is estimated[4].

In texture feature extraction the Co-occurrence matrix for feature extraction and Principle component analysis(PCA) for the reduction of the dimension is used. The images are first preprocessed, which mainly includes: get gray image, deal with the noise, and extract the interested area. Secondly, the image texture features are extracted in the area of interest by gray level co-occurrence matrix. Then dimension reduction is implemented by principal component analysis (PCA) to get the high correlation of 4-D feature vector. Tamura image texture features are extracted by Tamura. Through feature fusion, an 8-dimensional feature vector is got. Finally, classification and prediction are obtained by using the SVM to get the estimation of crowd density[5].

B. Crowd motion detection

The motion of the crowd can be detected by using Background Subtraction Method. As name suggested, the Background Subtraction is method of separate out the foreground object from background object in a sequence of video frames. A foreground object can be described as an object of attention which help in reducing the amount of data to be processed as well as provide important information to the task under
consideration. Often, the foreground object can be thought of as a coherently moving object in a scene. The Background subtraction is a class of techniques for segmenting out objects of interest in a scene for applications such as surveillance. There are many challenges in developing a good background subtraction algorithm. First, it must be robust against changes in illumination. Second, it should avoid detecting non-stationary background objects and shadow cast by moving objects. A good background model should also react quickly to changes in background and adapt itself to accommodate changes occurring in the background such as moving of a stationary chair from one place to another. It should also have a good foreground detection rate and the processing time for background subtraction should be real-time.

The problem in the background detection is that there are sudden changes in background i.e. the image is not fixed such as gradual changes, sudden changes, motion changes.

Background subtraction involve calculating a reference image, subtracting each new frame from this image and thresholding the result. What result is a binary segmentation of the image which highlights regions of non-stationary objects. The simplest form of the reference image is a time-averaged background image.

1. Estimate the background for time t.
2. Subtract the estimated background from the input frame.
3. Apply a threshold, Th, to the absolute difference to get the foreground mask.

After the foreground objects have been segmented, they are separated or identified using some algorithm to detect these regions. One of the ways is by using the region identification algorithm of computer vision where each separated objects are labeled differently to be able to distinguish them [6].

C. Crowd tracking

Object tracking is used for much application such as motion-based detection, automatic surveillance, video indexing and rescue, human computer communication and traffic monitor etc. Object tracking can be classified into several steps, such as object illustration, characteristic selection for tracking, object detection, background subtraction, and object segmentation. Object tracking is one of the most significant mechanism in a broad range of application in computer visualization, such as surveillance, human computer communication, and medical imaging. When persons are being tracked, the in order is compute at the level of every person. But when we chat about crowd tracking, the centre of attention is on the movement of a crowd as a set of tiny element whose arrangement change continuously, instead of tracking the same entity throughout the video. A relevance of crowd tracking is the use of crowd tracking to create models of crowd activities and to detect abnormal behaviors at the crowd level rather than at the individual level[7].

Aim of Blob detection method is detecting regions in a digital image that differ in properties, such as brightness or color, compared to surrounding regions. BLOB stands for Binary Large Object and refers to a group of connected pixels in a binary image. The term “Large” indicates that only objects of a certain size are of interest and that “small” binary objects are usually noise.

Also the tracking is done by the human head, which is difficult when the task is to monitor and manage large crowds in gathering areas such as airports and train stations. This technique make impossible to track large numbers of people in very crowded scenes in which the majority of the scene is in motion and most of the humans' bodies are partially or fully occupy. Under these conditions, we believe that the human head is the only body part that can be robustly detected and tracked, so a method for tracking pedestrians by detecting and tracking their heads rather than their full bodies.

Figure 5: Human Tracking using Blob Detection

Figure 6: Model of Human Tracking

In this system assume a single static camera placed at certain height that the heads of people traversing the scene can be observed. For initial detection use a standard AdaBoost Detector, and for tracking use a particle filter for each head that incorporates a simple motion model and a color histogram-based appearance model[9].
crowd behavior recognition is an important in video surveillance for public places. Normal behaviors are usually characterized by regular motion direction, low (walking) speed, stopping, etc. Abnormal behaviors include opposite movement in crowd, bifurcation, deviation, fighting, and are characterized by corresponding motion attributes[10]. Therefore, crowd behavior can be inferred from crowd motion information, carried by the long-term and short-term motion vectors. The motion attributes to be used in the crowd behavior model: principal directions, speed, and crowd mobility. Principal directions are simply computed from the direction histogram of crowd motion vectors. Suppose the majority of crowd vectors are aligned in one direction, i.e., there is one principal direction. The secondary principal direction, if existing, should involve only a small number of vectors. For instance, abnormal crowd events can be declared if there are multiple principal directions. The speed is directly related to the length of crowd vector. In video surveillance, the camera is usually fixed. Therefore, if camera is calibrated and the normal walking speed is estimated and the video frame rate is known, the vector length corresponding to the normal crowd speed can be determined. Crowd mobility is determined from the number of valid features. Within a reasonable short interval, the texture of a crowd image does not change much so the number of features should stay constant. This attribute can thus be used to determine abnormal events if the number of valid features changes drastically.

III. People Counting Model

In public venues, crowd size is a key indicator of crowd safety and stability. With the accelerated process of urbanization and socioeconomic, people gather in public more and more, which results in many crowded public emergencies. For reducing the number of crowd accidents, estimating the people number and crowd density should be very pivotal. As the crowd counting and crowd density are essential descriptors of all the status of crowd, it is very useful for us to learn the distribution of crowd and find out the tendency of abnormal behaviors, if we know more about the people number and density of the crowd. So crowd counting and crowd density are very useful information for security departments.

1. Models training: Divide the camera sample data into two categories of low-density crowd image and high density crowd image. Then, get the low-density crowd estimation model and high-density crowd estimation model by the model training system respectively. The model parameters are stored in the database o model, and associated with the GIS system.

2) Crowd counting estimation: Set the threshold to distinguish the high-density crowd and low-density crowd. Adopt low-density estimation model to estimate the people number of the crowd, if it is greater than the threshold, then choose the high-density estimation model.

3) Storage and management of crowd estimation results: After crowd estimation complete, the results were saved in the GIS database for the historical data query and analysis.

4) Spatial-temporal analysis and visualization: According to the estimation results, the number and density levels of current monitoring areas can be displayed on the map. And each monitoring region’s crowd flow were displayed in the form of statistical curve at the same time[12].

Another way of counting the people is by the blob detection, in which Median Filter is used to segment the foreground from the background and blob analysis is done to count the people in the current frame. Video from the camera is recorded frame by frame. In the first step background image is calculated and subtracted from all the images to get the foreground images. In the next stage, blobs of particular size are identified to estimate the number of people in the crowd[13].

IV. Benefits and Drawbacks Of Video Surveillance

A. Benefits

1. Video evidence: High-resolution footage and sound recordings capture everything that happens on property. Even if the main purpose is to catch criminals who attempt to steal from your inventory or vandalize the property, we could end up capturing far more than intended.

2. Solid support for insurance claims: Recorded footage of a theft or vandalism should be enough to validate the insurance claim. Plus, even if you never use footage in court, for investigations, or to make insurance claims, the lower insurance premiums that result from installing security cameras help your investment pay for itself.

3. Aid for your alarm system: One of the best ways to realize all the benefits of security cameras is to combine security cameras with a robust business alarm system. When an alarm sounds, a thief may panic and give something away. With this
combination, we can stop wrongdoers in the act and also capture footage of their attempts for court use.

B. Drawbacks

1. Encourage complacency : Security cameras should be one part of a more comprehensive business security system complete with alarms, electronic access control panels and environmental hazard detectors (smoke alarms, carbon monoxide detectors and sprinkler systems). You need additional security measures to ensure complete protection for your business.

2. Can’t stop robberies in progress : Cameras capture footage so you can receive justice in court, but they can’t stop a robbery in action. They don’t alert police like an alarm system does. That means, even with the perpetrator behind bars, you’ll still have intangible losses to deal with, such as the goodwill lost when customers no longer feel safe, and the time wasted in court, making insurance claims and reordering stolen inventory.

3. Can be useless in police investigations : Most thieves will do everything they can to hide their identity while doing a misdeed. If the perpetrator is masked and hooded, the footage picked up by a security camera can prove worthless in an investigation.

V. CONCLUSION

Video surveillance and crowd analysis are the most important concept for understanding the behavior especially in analyzing to detect abnormal behavior in crowd. We have to cope various obstacles to tackle the problem in order to make the detecting abnormal in crowd successfully. The obstacles maybe come from the different characteristic (involving different pose, position, velocity and density), changes of motion over time in a real video sequence, the occlusion surrounding people in crowd, illumination changes (transition from day to night, shadow of background images and non static background like leaves blown by the wind), and multiple input channel with the amount of number of camera where put many sides. However, with the variant analysis apply in a crowd analysis, these obstacles could be solved.

REFERENCES


[9] Irshad Ali and Matthew N. Dailey, "Multiple Human Tracking in High-Density Crowds".