Chapter 94
Early Recognition of Suspicious Activity for Crime Prevention

M. Kalaiselvi Geetha
Annamalai University, India

J. Arunnehru
Annamalai University, India

A. Geetha
Annamalai University, India

ABSTRACT

Automatic identification and early prediction of suspicious human activities are of significant importance in video surveillance research. By recognizing and predicting a criminal activity at an early stage, regrettable incidents can be avoided. Initially, an action recognition framework is developed for identifying the suspicious actions using interest point based 2D and 3D features and transform based approaches. This is subsequently followed by a novel approach for predicting the suspicious actions for crime prevention in real-world scenario. The prediction problem is formulated probabilistically and a novel approach that employs the mixture models for prediction is introduced. The developed system yields promising results for predicting the actions in real-time.

1. INTRODUCTION

Video surveillance is attracting much of the researchers' attention since it is a crucial tool for protecting people and public property and finds most promising application in computer vision. The recent acts of terrorism have necessitated the vital need for well-organized surveillance of suspicious human behaviors at important public places. Suspicious behavior in general incorporates an event that builds a disbelief or mistrust. These events have to be reported for further critical examinations. Currently the surveillance cameras are used extensively and the purpose of these cameras are primarily for security monitoring and identifying illegal actions and events. It is unfeasible for a human supervisor to inspect a huge volume of video recordings, since security personnel are also required to manage other tasks, such

DOI: 10.4018/978-1-5225-5204-8.ch094

as access control, handling emergency calls, following up on fire alarms, radio communications control, etc. However, this is still a novel technology with many practical restrictions. How can one differentiate among an individual running because he/she is delayed or because he/she has just committed a crime? Under such scenario, activity detection alone is not sufficed; prevention of such crime becomes necessary. Current researches on video analysis for surveillance system are post investigation methods and necessitate real-time processing. For example, recognizing missing objects after they have been stolen is of no use. Hence, it is enviable to predict the intension of a person, before the complete happening of the suspicious event. This prediction will possibly prevent a criminal activity, thus providing enough time to the security person to react upon the critical condition to ensure public safety. Hence an automatic activity recognition methodology is necessary.

A greater part of the present surveillance systems only record the events and act as a post event investigation tool which is not preferred. Other systems still submissively observe the surroundings and raise the alarm once a suspicious action is detected. Such an alarm will not help to prevent an unwanted event. Predicting a person's action before it is executed has an ample choice of applications in autonomous robots, surveillance and health care. While comparing the action recognition using full length video, action prediction with unfinished video observations requires identifying the well-grained factors intrinsic to the existing observations that would direct to the future action. For example, a person with stretched open arms indicates that he/she is going to hug. Although, traditional models like hidden markov models are used to approximate the prediction problem, they are found to be inappropriate for the sparse discontinuous features of the video. Thus it is indispensable to develop a novel prediction methodology for recognizing/predicting an incomplete activity from a video.

Activity recognition can be sensor-based or vision-based and this chapter addresses the problem of activity recognition using the later approach. In vision-based activity recognition, the computational procedure is done at four steps viz. human detection, tracking, activity recognition and then a complex activity evaluation.

Latest enhancement in the digital media makes people to produce their own digital video information. This creates the problem of categorizing the new video sequences based on the action categories. Apparently classifying this information for future reference manually is challenging and requires automated techniques. An effective approach to the detection of small objects is seen in (Hsieh, Han, Wu, Chuang & Fan, 2006) by employing watershed-based transformation. The proposed detection system comprises of two main modules, locating region of interest (ROI) and contour extraction. An image differencing technique is first employed on two neighboring image frames to produce rough candidate objects appearing in the images. A novel framework for contour based object detection from cluttered environments is discussed in (Lu, Adluru, Ling, Zhu & Latecki, 2010). The contour model for a class of objects is hierarchically decomposed into portions. Then, they are combined into part bundles, where a part bundle can contain overlapping portions. A background modeling that is appropriate to any spatio-temporal non-parametric moving object recognition strategy is proposed in (Cuevas & García, 2013). Through a proficient and strong method to dynamically estimate the bandwidth of the kernels used in the modeling, the efficiency of previous approaches are improved. Furthermore, by adding a novel mechanism to selectively revise the background model, the number of misdetections is decreased significantly, thus attaining improved performance. (Grabner, Leistner & Bischof, 2008) built a tracking method that utilizes variation in the appearance of the object due to lighting and rotation variations. To track the object using selected distinguishable low-level color features is suggested by (Collins, Liu & Leordeanu, 2005).