A NOVEL FRAMEWORK FOR MINING SUSPICIOUS ACTIVITIES IN SURVEILLANCE VIDEOS

[1]VIJAYAKUMAR VELUSAMY, [2]MINI T V

[1]Professor in Computer Science, S.N.R. Sons College, Coimbatore, Tamilnadu, India.
[2] Department of Computer Science, Sacred Heart College, Chalakudy, Kerala, India.

Abstract - The advancement of video surveillance systems have recently captured the interest of both research and industrial worlds due to the growing techniques of analytics in the security domain. The video analytics extracts useful information for protecting unlawful events and providing high level security. The application of data mining techniques detects of unusual or suspicious events in prerecorded and real time video-surveillance sequences. There are lot of solution framework exists those are tend to be highly domain specific. This paper presents a generic framework for mining the suspicious events in surveillance video. It also discusses about the efficient video pre-processing techniques, moving object detection and segmentation, object tracking and research issues and applications of intelligent video surveillance system.

Keywords - Suspicious event detection; Moving object segmentation; Object tracking; Video Analysis; Event mining;

I. INTRODUCTION

Video Surveillance Systems are widespread and common in many environments such as in indoor commercial places (hotels, supermarkets, banks) and indoor private areas (schools, colleges, industries) outdoor commercial locations (bus stands, railway stations and road traffic) monitoring the activities, or protecting for detecting unlawful activity, and providing high level security. The video event detection and analysis are an important activity in video surveillance systems. It can be manual, semi-autonomous or fully-autonomous. Manual video analysis involves analysis of the video content by a human. Semi-autonomous analysis encompasses some form of video processing but with significant human intervention. A fully-autonomous performs the entire task automatically without manual intervention. Generally, analysis of video requires combination of image processing techniques, temporal feature analysis, artificial intelligence and machine learning techniques. Image processing techniques are used to provide low level image (frame) features which extract visual cues and primitive events. Artificial intelligence techniques are used to provide expert decisions for the detection of more complex and abstract behavior pattern in a video sequence with considering temporal features. Machine learning techniques and algorithms are used to detect the objects, tracking and recognition tasks in the complex environment.

Data mining techniques is one of the efficient analytical tools for investigating data. It analyzes the data from many different dimensions, categorize it and summarize the relationships identified. There are various tools and product available in market for mining the video database. These tools and techniques focuses on extracting semantic information, implicit patterns and knowledge from video data and find correlations and patterns previously unknown from them [Vijayakumar and Nedunchezhian, 2012]. There have been some efforts about video data mining for movies, medical videos, traffic videos and surveillance videos. The main objectives of surveillance video mining are to obtain a description of what is happening in a monitored area such as real time automatic object identification, moving object behavior, object movement pattern recognition, modeling and detection of unusual/suspicious/abnormal/interesting events and recognition of event pattern. Mining the above semantic information in the surveillance video is one of the important benefits in developing security applications. An appropriate action can be taken based on the interpretation. It is very difficult to develop a general framework for mining events from surveillance video due to complexity of the video, method of feature extraction, noise removal methods and selection of suitable algorithms for mining. Generally, mining the video surveillance requires two successive steps such as surveillance video frame (image) analysis and mining the contents.

In this paper, a novel generic framework is proposed to automatically detect events, classify them into the right category and generates the appropriate action that can be used for further. The detail study of surveillance video data analysis is given in section 2. A review of mining suspicious events is presented in section 3. In section 4, a generic framework is proposed for mining suspicious events. In section 5, the research challenges and issues of suspicious event mining is presented. In section 6, summary of the paper is presented and future directions is discussed.

II. SURVEILLANCE VIDEO ANALYSIS

Video analysis detects and determines temporal and spatial events to automatically analyze video. It focuses on understanding and/or extracting features from the video database. Different types of functionality can be implemented in video analysis such as video motion detection, video tracking and behavior analysis or other forms of situation awareness.

Due to the unstructured nature of video data, the retrieving of video semantic information directly is not possible. There is a gap between the low-level video features such as color, texture, shape and high-level semantic concepts or events [Vijayakumar and Nedunchezhian, 2012]. Generally, five key steps involved in video analysis such as video pre-processing, detection of interesting moving objects, feature extraction and moving object segmentation, tracking of such objects from frame to frame and investigation of object to recognize their behavior/events [Rupesh Kumar Rout, 2013].

2.1. Video Preprocessing

Surveillance video analysis can be complex due to the noises in video frames; complexity of identifying the object motion due to camera motion, partial and full object occlusions, complex object shapes, scene illumination changes and real-time processing requirements [Makandar et al., 2015]. To overcome these problems, it is very essential to use efficient techniques on video processing system. In video analysis process, the preprocessing is done with extracting the Key Frame from the video sequence. Since a large number of adjacent frames are likely to be similar, one or more key frames can be extracted from the shot depending on the complexity of the content of the shot. The key frames extracted from the video streams are treated as images [Vijayakumar and Nedunchezhian, 2012]. The video frame (image) represents the salient visual contents of a shot. Many Techniques were introduced in the Key Frame Extraction [Vijayakumar and Nedunchezhian, 2011] such as Histogram Difference, Markov chain Monte Carlo (MCMC) technique, Intra-frame and inter-frame motion histogram analysis, Clustered based analysis and Generalized Gaussian Density.

The frame noise reduces visual quality and the efficiency of subsequent processing tasks. The classical image de-noising algorithms which are used to reduce noise are divided into two categories: the first process every frame independently as still image. The other type makes use of correlation between adjacent frames to get better effect. There are several types of noises which degrade the quality of video processing such as median noise, mean noise, impulse noise, Gaussian noise and bilateral noise etc., In pre-processing, the noise is reduced from video using noise reduction method in order to get the desired result. Noise may include out of focus subject, blur subject, motion blur due to camera motion while recording subject. The background noise is reduced using filter technique. Depending on the estimated noise, noise levels and type of noise the reduction method can be selected for the noise reduction.

The median filter technique is a non-linear filtering and it preserves edges while removing noise. A bilateral filter preserves the edge and removes the noise. It is a non-linear filter. The intensity value at each pixel in the image is replaced by a weighted average of intensity values from nearby pixels. The average/mean filter performs spatial filtering on each individual pixel in an image using the grey level values in a square or rectangular window surrounding each pixel [Hoshyar et al., 2014]. The goal of morphological image processing is remove imperfections by keeping the structure of video

2.2. Moving Object Detection and Segmentation

It is a way of defining an activity in a scene by analyzing the differences that occur in a series of images. Video Objects are the set of semantic objects contained within the video object like car, bike and human. Object detection in videos involves confirming the existence of an object in video frame sequences and discovering it exactly for recognition. Detecting moving objects is a major significance in video object detection and tracking. The pixel matching or frame referencing is use for object detection. Any change between frames is regarded as ‘detection’. The background subtraction, statistical methods, temporal differencing and optical flow are the frequently used moving object detection techniques [Shaikh et al., 2014]. These are use low level information about each pixel. At present, most segmentation methods use either temporal or spatial information in the image sequence.

Background subtraction identifies moving areas by taking the transformation between the present image and the reference background image in a pixel-by-pixel method. The image is usually captured by a stationary camera, it is easier to detect a still background than moving object. The moving regions are extracted by temporal differencing such as the pixel differences between the consecutive frames in a sequences. The flow vectors of moving objects are used in the optical-flow-based motion segmentation to detect moving regions.

In Computer Vision, moving object segmentation refers to the process of identifying and partitioning the meaningful regions present in video sequences such as vehicles and humans. Accurate and real-time moving object segmentation will greatly improve the performance of object tracking, activity analysis and high-level event understanding [Solana-Cipres et al., 2009].

Static foreground objects are entering into the scene, reach a given position and then stop their motion. These objects usually re-start their motion after some period of time. Examples are cars in parking sequences, people in smart rooms, etc. Abandoned luggage is also static foreground that is interesting in some applications. The possibility to distinguish between moving and static objects is based on the fact that when an object remains static, its pixels are all assigned to the same moving foreground model for a certain period of time, and thus the counter will increase. On the contrary, when an object is moving, its foreground pixel model needs to be re-initialized often in successive frames [Gallego and Pardas, 2008].

The moving object detection and segmentation algorithms can identify with accuracy the border of the regions, but the most of algorithms cannot work in real-time. Fast algorithms work directly on compressed video to segment moving objects.

2.3. Video Feature Extraction

It refers to the process of extracting the useful characteristics from a frame in a video. It is important to detect and track an objects in a surveillance video. Objects can be easily distinguished in the feature space with most desirable visual features such as Color, Edges, Optical Flow and Texture. This step extract, find out, match the feature or object in one frame of the video (e.g. find out the ball in a frame). It checks object location and moving objects in sequential frames and a rectangle box is marked around the object using its dimensions to track the objects.

Shape is an important visual feature for image content description. Shape descriptors can be classified as region based and contour-based methods. Region-based methods use shape description and contour-based methods uses contour of an object [Choras, 2008].

Motion features are useful to identify the semantic concepts and it can be classified into two categories such as camera-based and object-based methods. The Video objects in camera focus have higher contrast and sharper edges compared to the background. This can be measured using specific features such as color, texture, motion. The moving object detection is essential step for object tracking and also requires corresponding motion parameters.

The individual features like trajectory, speed, gait and specific object features are used to detect the usual or suspicious person behavior in the preventing unlawful activities. The trajectory features of moving objects were extracted and represented as a feature vector. Behavior understanding includes the analysis of motion patterns, and the production of high-level description of actions and interactions.

In the traffic management, the object features such as vehicle features are used in the road traffic statistics analysis. The person, the vehicle features and the time of occurrence are used to detect the anomaly behavior like theft detection in parking lots, hypermarkets. The image of vehicle number plates features are used recognize the vehicle number plate for detection of stolen vehicles or vehicles with illegitimate plates (that may serve for unlawful activities).

2.4. Object Tracking and Classification

After motion and object detection of interesting objects, surveillance systems generally track moving objects from one frame to another in an image sequence. The object detection aims to locate and segment interesting objects. Then, such objects can be tracked from frame to frame, and the tracks can be analyzed to recognize object behavior. Tracking over time typically involves matching objects in consecutive frames using features such as points, lines or blobs [Teddy ko, 2008]. The purpose of tracking is to determine the spatial temporal information of each target present in the scene.

Selecting the right features play a critical role in tracking. It is closely related to the object representation. For example, color is used as a feature for histogram-based appearance representations, while for contour-based representation object edges are usually used as feature. In general, many tracking algorithms use a combination of these features.

Object Tracking methods are divided into four major categories: region-based tracking, active-contour-based tracking, feature based tracking, model-based tracking and hybrid tracking. Region-based tracking algorithms track objects according to variations of the image regions corresponding to the moving objects. These algorithms cannot satisfy the requirements for surveillance against a cluttered background or with multiple moving objects. Active contour-based tracking algorithms track objects by representing their outlines as bounding contours and updating these contours dynamically in successive frames. Feature-based tracking algorithms can be categorized into global feature-based algorithms, local feature-based algorithms, and dependence-graph-based algorithms. Model-based tracking algorithms track objects by matching projected object models, produced with prior knowledge, to image data.

The surveillance systems tracks of static foreground regions based on the tracking of abandoned, stolen objects, parked vehicles. A tracking pipeline developed to investigate different combinations of foreground extraction, feature extraction and motion correspondence algorithms [Karasulu and Korukoglu, 2013]. In foreground extraction they explored applications of background subtraction techniques and salient region extraction. They examined the application of salient region detection to extracting moving objects and focused on the use of Kadir and Brady's Scale-Saliency algorithm and Lowe's SIFT key points. In feature extraction, they extract features such as the centroid, the size, and the average pixel intensity of each moving object. These are then used in tracking algorithms such as Kalman filters to track the objects from one frame to the next.

To further tracking and analyze the behaviors of objects, it is essential to correctly classify moving objects. The moving object classification task interacts with the tracker in each frame, voting for the class of each detected target. Currently, there are two major approaches towards moving object classification which are shape-based and motion-based methods [Weiming Hu et al., 2004]. Shape-based methods make use of the objects' 2D spatial information whereas motion-based methods use temporally tracked features of objects for the classification solution. Shape-based classification uses points, boxes, silhouettes and blobs for classifying moving objects. Temporal consistency constraints are considered to make classification results more precise.

2.5. Behavior Analysis

Behavior understanding recognizes the motions and creates of advanced description about activities. It also investigates the interactions between or among objects such as detection and understanding of human behavior around a parked vehicle.

Moving object detection and tracking carries information about the spatio-temporal relationships between objects. This information, enabling us to identify normal patterns, object behaviors and detect unusual events. The detection and tracking of moving objects is a task which must be performed accurately and robustly to minimize false alarms and missed positives, and in real-time to enable corrective action. Generally, the low level sampling features are used to detect the activity in the video surveillance systems [Tziakosa et al., 2010].

Three different kinds of behaviors are identified in the security domain such as normal, unusual and abnormal. The surveillance area focuses unusual behaviors, such as a person running in a hostel entrance and violation of restricted areas, e.g. a person crossing a railway.

Brand and Vettnaker [2000] developed a system that learns patterns of behavior by training a HMM. Anomalous behavior, such as a car driving in the wrong lane or turning left from the right hand lane, is then detected from a video camera mounted above an intersection. This system is also applied in an office environment to detect unusual behavior (such as falling a sleep, or standing at the window, apparently).

III. MINING THE SUSPIOUS ACTIVITES IN SURVEILLANCE VIDEO

Due to the increasing demands of intelligence, there is a need of video data mining techniques used for finding implicit, useful and knowledge from a large number of video data. They can help us understand video solutions automatically, improve intelligence of surveillance applications and make decisions. An intelligent surveillance system could provide not only the recording function but also the detection of abnormal activities. A key task in mining surveillance video is the detection and tracking of moving objects and analysis of object behavior, such as tracking the people and vehicles, through the video frames and identifying the activity of a human or vehicle . It consist sequence of steps such as video analysis and video mining. The first step is analysis the content and structure of videos, extraction of features of moving object, spatial or temporal correlations among those features. In the second step, discovers patterns of video structure, objects activity, video events, etc. from wide amounts of video data without little assumption concerned with their contents.

Lot of knowledge can be mined from the surveillance video database, such as, Classification of an object behavior, Spatio-temporal pattern or sequence classification, Characterization of class, or event e.g. for additional learning or for traffic analysis, Clustering and outliners detection and etc.,

Gowsikhaa et al., [2012] proposed a method to detect suspicious activities in the examination hall such as object exchange, entry of new person/unauthorized person, peeping into other’s answer sheet and person exchange from the video captured by a surveillance camera during examinations.

M. Brand and V. Kettnaker[2000]presented a method of video mining for identifying the abnormal and meaningful events. A background subtraction algorithm was utilized to get the binary mask of moving objects. The motion of every frame calculated to segment the sequence of surveillance video. Then a mixture of hidden Markov models using the expectation-maximization scheme fitted to the motion data with some probability to identity the similar segments. Finally, abnormal events and meaningful patterns are mined.

Dai et al., [2006] proposed a method for abnormal event detection in a surveillance video. A surveillance video data mining approach is proposed to discover similar video segments from surveillance video through a probabilistic model. The details moving vehicles are obtained by background subtraction method then tracked the detected vehicles with fast normalized cross-correlation method and extracted the interested video information, such as vehicle I/O time, vehicle type, vehicle color, etc. The useful information for decision can be gotten by video querying and statistical analysis based on the structured data.

Oh, and Bandi [2002] proposed a framework to perform the tasks such as temporal segmentation of video sequences, motion feature extraction, and clustering of segments. They considered color features for more clustering and indexing.

Thuraisingham [2004] presented the method for mining for various types of threats in surveillance video mining. They also discussed the applications of video data mining for national security and cyber security.

Mining the surveillance video might handle in the various application scenarios such as the parking lots are examined for crime prevention and detection. The activities of elderly and infirm people are observed for early alarms. The effectiveness of medical treatments is measured. National borders are monitored for measuring flow of refugees.

IV. GENERIC FRAMEWORK FOR MINING THE SUSPICIOUS ACTIVITIES

Video Analytics analyzes live or recorded video and generates data for pre-emptive action or data mining purposes. It is generally regarded as superior to video event detection and behavior analysis as it identifies items of interest, eliminates the cause of many false alarms, identifies suspicious events, abnormal activities and can provide more useful information. In this chapter, a generic framework for mining suspicious activity surveillance system is proposed based on the motion features shown below in the figure.1.

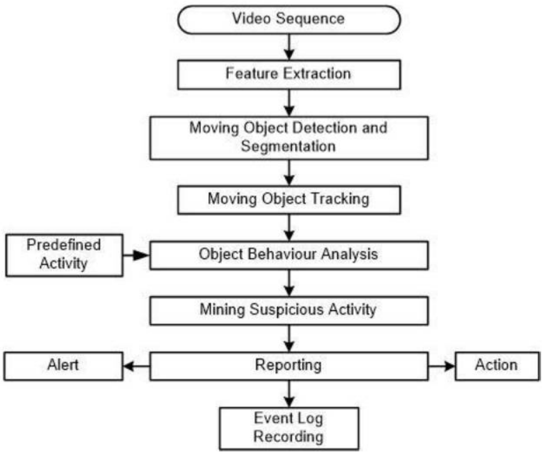

Fig. 1. A Generic Framework for Mining Suspicious Activity

In the video preprocessing step, the key frames are extracted using color histogram techniques. The frame noises are removed by applying the efficient Gaussian filter noise reduction technique. In the second stage, Moving objects are tracked to extract motion information with a frame difference method is applied to detect the moving object. Next particle filtering techniques is applied track the objects. K-Means clustering is applied to identify and cluster the suspicious activities. Finally, an alert is generated for the suspicious activity.

V. APPLICATIONS AND RESEARCH CHALLENGES

There is a lot of applications of surveillance video mining in various domain such as government, enterprise, military but not restricted to [Saravanan and Kanimozhi, 2010]. Automatic analysis for surveillance video such as detecting patterns; recognition of suspicious people in large crowds; detection of traffic patterns; monitoring of surveillance in theft or fire protection, care of bedridden patients and young children are important applications of video mining.

The main issues in object tracking are detection of moving objects and analysis of activity. It can be difficult for several reasons like motion of the camera, partial and full object occlusions, complex object shapes, scene illumination changes, and real-time processing requirements, changes in illumination of a scene, objects such as waving trees, objects that come to a stop and move again such as vehicles at a traffic light, etc . Surveillance generally demands that objects are tracked over long periods of time, and in varying conditions. This raises difficulties such as tracking in very different lighting conditions (possibly day and night), across a cluttered and dynamic background, and in the presence of shadows. The types of features chosen for analysis that can adequately capture the spatial and temporal signatures of the different behaviors in the scene. The noise in the features of interest and still ensure robustness. Video object motion analysis, representation of effective video data model, extracting the appropriate video features, and selecting the video data-mining algorithms are the challenges in the video analysis. The surveillance area can be extended with multiple cameras and the vision information can overcome occlusion. Single camera tracking produces ambiguity due to occlusion or depth and it may be eliminated from another view.

CONCLUSION

This paper presents the generic framework for mining suspicious activity which is the major challenge in the video data mining. It also reviews and exploits general strategies of stages involved in mining of video surveillance. This paper analyzes the challenges object tracking, motion analysis, behavior analysis, and mining of suspicious events.

REFERENCES

[1] Alper Yilmaz and Mubarak Shah, Object Tracking: A Survey, ACM Computing Surveys (CSUR) Surveys Homepage archive, Volume 38, Issue 4, No. 13, 2006.

[2] Azadeh Noori Hoshyar , Adel Al-Jumailya , Afsaneh Noori Hoshyar, Comparing the Performance of Various Filters on Skin Cancer Images, International Conference on Robot PRIDE 2013-2014 - Medical and Rehabilitation Robotics and Instrumentation, ConfPRIDE 2013-2014, Procedia Computer Science 42 (2014), pp. 32 – 37.

[3] Aziz Makandar , Daneshwari Mulimani, and Mahantesh Jevoor, Preprocessing Step- Review of Key Frame Extraction Techniques for Object Detection in video, International Journal of Current Engineering and Technology, Vol.5, No.3, June 2015.

[4] Bhavani Thuraisingham, Data Mining for Security Applications, Proceedings Of International Conference on Machine Learning and Applications, 16-18 Dec. 2004, pp. 3 – 4.

[5] Brand . M., and Kettnaker, V., Discovery and segmentation of activities in video. IEEE Trans. Pattern Analysis and Machine Intelligence, Volume: 22, No:8, 2000, pp:844–851.

[6] Gallego J, and Pardàs, M. Segmentation and Tracking of Static and Moving Objects in Video Surveillance Scenarios, In: IEEE International Conference on Image Processing. 2008 pp. 2716–2719.

[7] Gowsikhaa D., Manjunath, and Abirami S., Suspicious Human Activity Detection from Surveillance Videos, International Journal on Internet and Distributed Computing Systems. Vol: 2 No: 2, 2012, pp. 141-149.

[8] Ioannis Tziakosa, Andrea Cavallaroa, Li-Qun Xub, Event Monitoring Via Local Motion Abnormality Detection in Non-Linear Subspace, Neurocomputing, 2010; pp:1881-1891.

[9] Karasulu, B., and Korukoglu, S., Moving Object Detection and Tracking in Videos, Performance EvaluationSoftware, SpringerBriefs in Computer Science, 2013, pp.7-30.

[10] Ke-Xue Dai, Guo-Hui Li, and Ya-Li Gan, A Probabilistic Model For Surveillance Video Mining, Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian, 2006, pp.1144-1149.

[11] Liang Wang , Weiming Hu, and Tieniu Tan, Recent Developments in Human Motion Analysis, http://www.cs.cmu.edu/~dgovinda/pdf/recog/PR%20.pdf.

[12] Lijie Liu, and Guoliang Fan, Combined Key-Frame Extraction and Object-Based Video Segmentation, IEEE Transactions on Circuits and Systems For Video Technology, Vol. 15, No. 7, July 2005, pp. 869- 884

[13] Oh, J., and Bandi, B., Multimedia data mining framework for raw video sequences. In: Proc. of ACM Third International Workshop on Multimedia. Data Mining (MDM/KDD2002), Edmonton, Alberta, Canada, July 2002.

[14] Prati, A., Vezzani,R., Fornaciari, M., and Cucchiara, R., Intelligent Video Surveillance as a Service in Intelligent Multimedia Surveillance: Current Trends and Research, 2013, pp.1-16.

[15] Rupesh Kumar Rout, A Survey on Object Detection and Tracking Algorithms, Thesis, National Institute of Technology Rourkela, June 2013.

[16] Ryszard S. Choras, Shape-based Image Retrieval, Proceedings of the 7th WSEAS International Conference on Signal Processing (SIP'08), Istanbul, Turkey, May 27-30, 2008, pp.99-104.

[17] Saravanan. D., and Kanimozhi,T., Information Retrieval: Using Video Data Mining Methods, International Journal of Computer Communication and Information System– Vol2. No1. July – Dec 2010.

[18] Soharab Hossain Shaikh, Khalid Saeed, and Nabendu Chaki Moving Object Detection Using Background Subtraction, Springer, Briefs in Computer Science, 2014, pp.15-23.

[19] Solana-Cipres , Fernandez-Escribano, Rodriguez-Benitez , Moreno-Garcia, and Jimenez-Linares, Real-time moving object segmentation in H.264 compressed domain based on approximate reasoning, International Journal of Approximate Reasoning, 2009, pp:99–11.

[20] Teddy ko., A survey on behavior analysis in video surveillance applications, 37th IEEE Applied Imagery Pattern Recognition Workshop, 2008, pp.1 – 8.

[21] Vijayakumar Velusamy and Nedunchezhian, R., A study on video data mining, International Journal of Multimedia Information Retrieval, 2012, pp:153–172.

[22] Vijayakumar Velusamy, and R.Nedunchezhian, A Novel Method for Super Imposed Text Extraction in a Sports Video, International Journal of Computer Applications, Volume 15, No.1, 2011, pp.1-6.

[23] Weiming Hu , Tieniu Tan, Liang Wang, and Steve Maybank, A survey on visual surveillance of object motion and behaviors, IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews , vol. 34, no. 3, 2004, pp. 334-352.

★★★