International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 04 Issue: 12 | Dec-2017 www.irjet.net p-ISSN: 2395-0072

INSPECTION OF SUSPICIOUS HUMAN ACTIVITY IN THE CROWDSOURCED AREAS CAPTURED IN SURVEILLANCE CAMERAS

P.Bhagya Divya[1], S.Shalini[2], R.Deepa[3] , Baddeli Sravya Reddy[4]

[1&4]*PG scholar, Department of Computer Science and Engineering,*
[2&3]*Assistant Professor, Department of Computer Science and Engineering,*
[1, 2,3&4] *Prince Dr.K.Vasudevan College of Engineering and Technology, Chennai, Tamil Nadu, India.*

Abstract - *The ultimate aim is to provide the indoor security using the CCTV camera. The CCTV Camera is a video camera that feeds or streams its image in real time; Webcams are known for their low manufacturing cost and their high flexibility, making them the lowest-cost form of video conversations and inefficient security issues. The system will detect suspicious person i.e. unauthorized entry in a restricted place in a video by using AMD algorithm and will start tracking once the user has specified a suspicious person by his/her on the display. The main purpose of background subtraction is to generate a reliable background model and thus significantly improve the detection of moving objects. Advanced Motion Detection (AMD) achieves complete detection of moving objects. A camera is been connected inside the monitoring room which produces alert messages on the account of any suspicious activity*.

Keywords: CCTV camera, Advanced Motion Detection, Background model, Suspicious activity, Webcams

1. INTRODUCTION

The region of the possibility and information program focuses on research in various security techniques to address problems in repeated aim detection applications. The goal of computerized surveillance system is to support the human operator in prospect investigation and event categorization by without human intervention detecting the objects and analyzing their actions using computer vision, techniques of pattern recognition and signal processing. This review addresses more than a few advancements made in these fields' while bringing out the detail that realizing a practical end to end surveillance system still remnants a hard job due to more than a few challenges faced in a real world situation. With the improvement in computing technology and now in-expensively and technically possible to adopt multi camera and multi-modal structure to gather the requirement of well-organized surveillance system in broad range of security applications like security guard for important buildings and surveillance in cities.

Visual surveillance has been an energetic study area in pattern analysis and machine intelligence, due to its vital position in helping surveillance intelligence and law enforcement agencies to battle alongside offense and crime actions. The objective of a visual surveillance system is to identify irregular object behaviors and to lift alarms when such behaviors are detected using the Advanced Motion Detection (AMD) algorithm.

$$M_d(x,y)=\begin{cases}1, if|I(x,y)-S(x,y)| > T\\0, if|I(x,y)-S(x,y)| \leq T\end{cases}$$

After moving objects are detected, it is necessary to categorize them into predefined categories, so that their movements behaviors can be suitably interpret in the background of their identities and their connections with the surface. Therefore, object categorization is a very important part in a complete visual surveillance system.

1.1 Related work

The SNV technology presents a novel framework for recognize human activities from video sequences captured by depth cameras. They extend the surface normal to polynomial by assembling local neighbouring excited surface normal's from a depth sequence to jointly characterize local motion and shape information. Then they propose a general scheme of Super Normal Vector (SNV) to cumulate the low-level polynomials into a discriminative representation [1]. Binary range-sample feature in depth is implemented. The goal is to engender front, activity, and back layers. Seeds for generating the two bounding planes to separate them are required. Joint points with depth less than zero can be naturally regarded as the front seed points. This is a very coarse operation, but is already sufficient in our feature construction. It has denoted them respectively as Cfront and Cback.[2]

A multi-part bag-of-poses approach is then defined, which permits the separate alignment of body parts through a nearest-neighbor classification. Experiments conducted on the Florence 3D Action dataset and the MSR Daily Activities dataset show promising results. This method has been evaluated on two samples: the Florence 3D Action Dataset and the Microsoft (MSR) Daily Activity 3D data set [4]. The curiosity in the capturing of human actions is motivated by the promise of many application, both offline and online. For gradually more large and complex datasets, manual labeling will become prohibitive. Automatic labeling using video subtitles and movie scripts is possible in some domains, but still requires manual verification. They discussed vision-based human action recognition in this survey but a multi-modal approach could improve detection in some domains

[5]. The temporal evolution of a modality appears to be well approximated by a sequence of temporal segments called onset, apex, and offset. The experimental results obtained show the following: 1) affective face and body displays are simultaneous but not strictly synchronous; 2) explicit detection of the temporal phases can improve the accuracy of affect recognition; 3) recognition from fused face and body modalities performs better than that from the face or the body modality alone; and 4) synchronized feature-level fusion achieves better performance than decision-level fusion.[6].

Trajectory captures the local motion information of the video. A dense representation guarantees a good coverage of foreground motion as well as of the surrounding context. Additionally, they present a descriptor based on Motion Boundary Histograms (MBH) which rely on differential optical flow. The MBH descriptor shows to consistently do better than other state-of-art descriptors, in particular on real-world videos that contain a significant amount of camera motion [9]. A filtering method is implemented to extract STIPs from depth videos (called DSTIP) that effectively suppress the noisy measurements. Further, they have built a novel Depth Cuboids Similarity Feature (DCSF) to describe the local 3D depth cuboids around the DSTIPs with an adaptable supporting size. Experimental evaluation shows that the proposed approach outperforms state of-art activity recognition algorithms on depth videos, and the framework is more widely applicable than existing approaches. They also give detailed comparisons with other features and analysis of choice of parameters as a guidance for applications[10].

2. EXISTING SYSTEM

Existing approaches requires the user to record a video at the faces and then process it to recognize them, although the picture taken by user may not be able to capture the image using the Depth cameras.

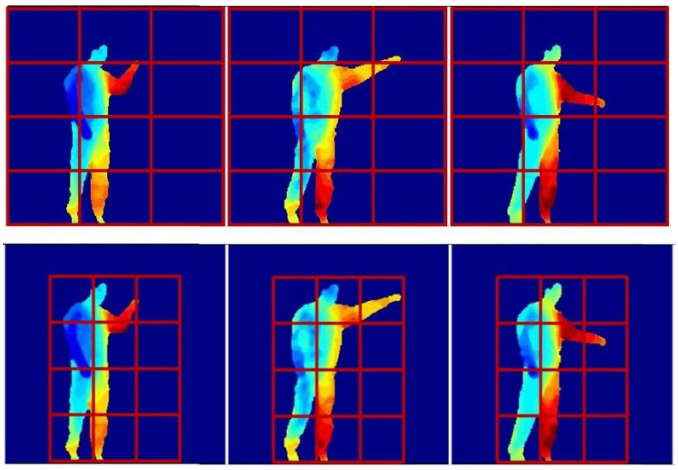

Fig 1: Detection of human actions by capturing the images using the Depth cameras [Ref 1]

The depth images are in the appearance of the shaded human figure as shown in the figure 1. They mainly concentrate on the study or action recognition of the humans which can be used for further examination on human civilization. The use the technique of Super Normal Vectors (SNV) and uses the implementation of polynormals.

$U=(u_1^T, \dots\dots, u_k^T)^T$ → Equation of polynormal from [1]

The images retrieved from the depth cameras images cannot be used for the identification of the human faces or some other unique identification. Existing researches has a major drawback of inefficiency in the case of online processing of videos for crime reduction.

3. PROPOSED SCHEME

The CCTV Camera is a video camera that feeds or streams its image in real time .The system will detect suspicious person i.e. unauthorized entry in a restricted place in a video by using AMD algorithm and will start tracking once the user has specified a suspicious person by his/her on the display. The main purpose of efficient background subtraction method is to generate a reliable background model and thus significantly improve the detection of moving objects.

Advanced Motion Detection (AMD) achieves complete detection of moving objects. A camera is been connected inside the monitoring room which produces alert messages on the account of any suspicious activity.

A. Background Modeling (BM)

Background subtraction, also known as foreground detection, is a technique in the field so image processing and computer vision where in an image's foreground is extracted for further processing (object acknowledgment, etc .).

$$M_d(x,y)=\begin{cases}1, if|I(x,y)-S(x,y)| > T\\ 0, if|I(x,y)-S(x,y)| \le T\end{cases}$$

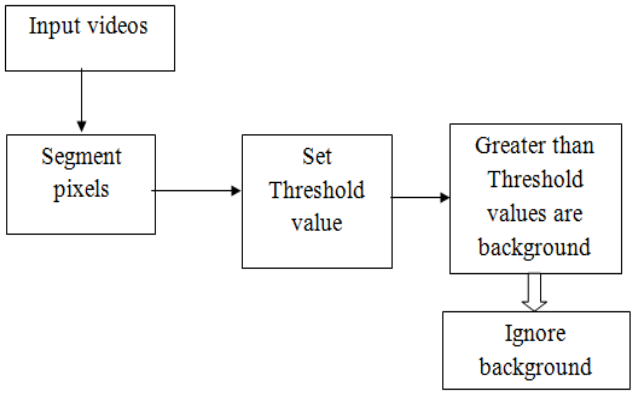

Fig 2: Advanced Background subtraction model

Generally an image's regions of interest are objects in its forefront. The sample image that is included in the video for detecting the moving objects cam be given as $S(x, y)$ and the image taken for motion detection can be given as M_d(x,y).Then I(x,y) is the input video frames.

The architecture of the suspicious human activity detection system is provided with seven modules that is included with an alarm trigger system.

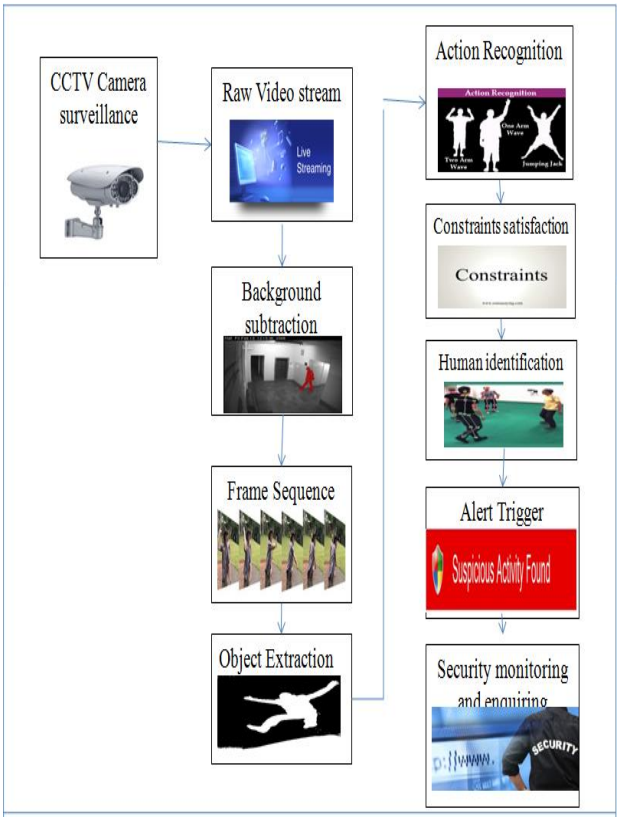

Fig 3: System Architecture for detecting the suspicious human activities

B. Frame Sequence

The video to frame detection can be done by using so many software's which are available in the market today however when we are using this software to get the frames from the video software will decide in the begin itself how many frames we need per second so which indicates that there will be a chances of missing the frames on which we are intent more, normally the number frames per second will be different for the different cameras.

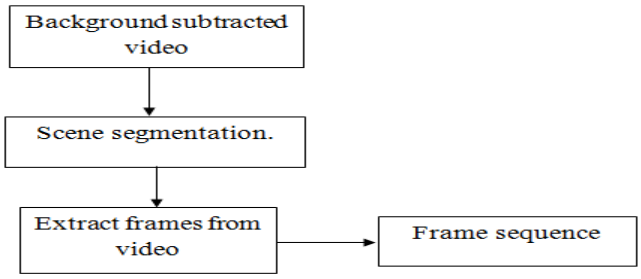

Fig 4: Conversion of raw videos into frames

C. Object Extraction

A new method of video object extraction is proposed to accurately obtain the object of interest from actively acquired videos.

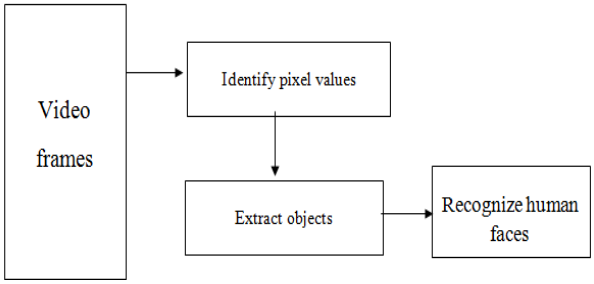

Fig 5: Extraction of images of the humans from the converted frame sequences

Traditional video object extraction techniques often operate under the assumption of standardized object motion and extract various parts of the video that are motion consistent as objects. In contrast, the proposed active video object extraction (AVOE) paradigm assumes that the object of interest is being actively tracked by a non-calibrated camera under general motion and classifies the possible actions of the camera that result in the 2D motion pattern as recovered from the image succession.

The result that is obtained as a output of the Phase I work is the identification of the human faces.It ca be given as the result:

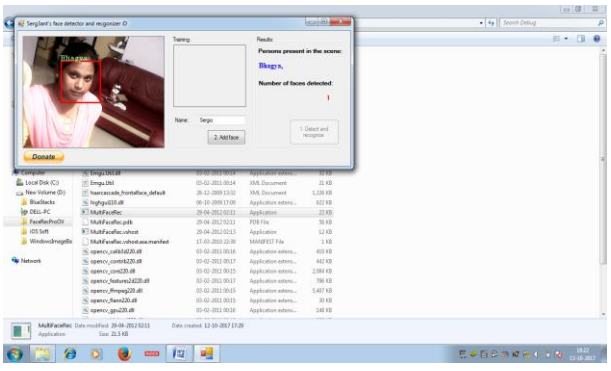

Fig 6: Phase I output screenshot

D. Detection of Suspicious activity

Detection of suspicious activity by video surveillances is highly effective. In previous decade monitoring of video by humans those are sited in front of screen of videos captured by either CCTV or any other cameras. Now we are going to automate this type of monitoring the best techniques which is used by most cases is image processing. Pattern analysis is a method of surveillance specifically used for documenting or understanding a subject's (or many subjects') behavior.
The system follows 3 main constraints such as , Height, Time and body movement. When the constrains are satisfied for the activities of a particular person, he will be considered as a doubtful person to be reported.

E. Advanced motion detection (AMD) algorithm

Algorithm 1: Computation of human action

Input: a frame images I
a coding operator C
a activity D = (dk)K
k=1
a set of space-time cells I = fs
Output: Action Detection
1 computes polynormals fpig from I
2 compute coefficients f_ig of fpig by R
3 for space-time cell i = 1 to jI j do
4 for visual word k = 1 to K do
5 uki
: = spatial average pooling and temporal
Max pooling of _i;k (pi ⍰ dk), where pi 2 vi
6 end
7 Ui :=
u1i
; : : : ;uKi
_
8 end

4. CONCLUSIONS

The system has presented a novel module that generated an accurate background with production of neither inefficient pixels nor artificial "ghost" trails. After a high quality background model was produced, the AT module eliminated the unnecessary examination of the entire background region and reduced the computational complexity for the consequent motion detection phase. The proposed object extraction module detected the pixel of moving objects within the triggered alert region to from the moving object mask. It also initiates the development of a system for suspicious human monitoring and study of their behaviors. Finally this algorithm works for Online (Real-time) video processing and its computational involvedness is low.

In future, the system can be used with the highly accessible storage service and it can also be implemented with hi-tech mode of capturing of videos in the surveillance areas.

REFERENCES

[1] "Super Normal Vector for Human Activity Recognition with Depth Cameras", Xiaodong Yang, Member, IEEE, and YingLi Tian, Senior Member, IEEE

[2] C. Lu, J. Jia, and C. Tang, "Range-Sample Depth Feature for Action Recognition", CVPR, 2014.

[3] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R.Moore, A. Kipman, and A. Blake, "Real-Time Pose Recognition in Parts from Single Depth Images", CVPR, 2011.

[4] L. Seidenari, V. Varano, S. Berretti, A. Bimbo, and P. Pala, "Recognizing Actions from Depth Cameras as Weakly Aligned Multi-Part Bag-of-Poses", CVPR Workshop on Human Activity Understanding from 3D Data, 2013.

[5] R. Poppe, "A Survey on Vision based Human Action Recognition", Image and Vision Computing, 2010.

[6] H. Gunes and M. Piccardi, "Automatic Temporal Segment Detection and Affect Recognition from Face and Body Display", IEEETrans. Systems, Man, and Cybernetics - Part B: Cybernetics, 2009.

[7] O. Oreifej and Z. Liu, "HON4D: Histogram of Oriented 4D Normals for Activity Recognition from Depth Sequences", CVPR, 2013.

[8] J. Luo, W. Wang, and H. Qi, "Group Sparsity and Geometry Constrained Dictionary Learning for Action Recognition from Depth Maps", ICCV, 2013.

[9] H.Wang, A. Klaser, C. Schmid, and C. Liu, "Dense Trajectories and Motion Boundary escriptors for Action Recognition", International Journal on Computer Vision, 2013.

[10] L. Xia and J. Aggarwal,"Spatio-Temporal Depth Cuboid Similarity Feature for Activity Recognition Using Depth Camera", CVPR, 2013.

[11] Advanced Motion Detectio(AMD) technique: http://ieeexplore.ieee.org/abstract/document/5605242/?reload=true

[12] video surveillance performance : https://www.videosurveillance.com/cctv-technology/cctv-video-management.asp

[13] Backgroung modelling strategy: https://docs.opencv.org/3.2.0/d1/dc5/tutorial_background_subtraction.html

[14] Human Face identification and recognition: https://facedetection.com/algorithms/

[15] Online video processing strategy: https://online.duke.edu/course/image-video-processing/