Asian Journal of Convergence in Technology **Volume V, Issue II**
ISSN No.:2350-1146, I.F-5.11

Anomaly Detection in Surveillance Video Using Pose Estimation

K.V.Prema
MTech in Network and InternetEnggineering
Department of Electronics andCommunication,
JSS Science and TechnologyUniversity
Email:kvprema2013@gmail.com

A Thyagarajamurthy
MTech (PhD),Prof
Department of Electronics and Communication
JSS Science and Technology University
Email:trm.sjce@gmail.com

Abstract—**Anomalous events are generally infrequent, sparse, and unpredictable. Detection of people and continuously monitoring them and base on the human activity detecting anomaly event is challenging task. In this paper, we focus on keypoints detection of a person. We judge the abnormal behavior of the person by detecting the motion of key-points of that person. In the starting frame we select a bounding box. If some number of key-points are moved out of the bounding box then system gives an alert message.**

Keywords—Bounding box,frame,keypoint

I. INTRODUCTION

Security problems are attracting increasingly more attention nowadays. Millions of surveillance cameras are placed to discover abnormal events, which creates the premise that an alarm can be released automatically when abnormal events occur. Therefore, developing an anomaly detection algorithm is of significant importance. There is an increasing demand for automatic methods for analyzing the vast quantities of surveillance video data generated continuously by closed- circuit television (CCTV) Systems. Digital image processing has been tremendous growth ever since it came in to existence. Digital image processing uses algorithm to process the image, after processing we are able to get the information which is relevant to our need.

Aim of this paper to have a potential security system. Security can be achieved by observing the people's movement in crowd or a person behavior in the crowd. By anomaly detection, we mean to detect motion patterns in the video that do not conform to the expected behavior. That is if some stranger who walks by the house feels abnormal then that behavior can be detected using the same cc camera which is backed by the imageprocessing.

The CCTV cameras are just recording the live scenarios ad transmitting it to a specified target. Though it is serving the purpose, it is not the end of the scope. The footage of the CCTV camera may serve as proof to a crime. Even in rare cases, the footages even help to identify the anomalies or misbehaviors and thus help the authority to prevent the happenings. With the advancement of the technology, thereis scope for everything. The field of image processing allows exploiting the limits and going much beyond.

Using image processing, one can develop algorithm that could detect any abnormalities in an image or a video automatically and conveniently. our aim is to bring in such a method that could be implemented in any surveillance camera so that there are higher chances of detecting anomalies, be it anything for example, Human actions like fighting, chasing and natural actions such as smoke and fire. The system not only detects the anomalies but also helps to prevent them by notifying the controlauthority.

II. LITERATURE REVIEW

To detect abnormalities in surveillance videos there are many approaches which are developed till now:

According to [1] there are three types of detection techniques: Vision based approaches, Physics inspired approaches and Physics simulated approaches. In [1] Physics simulated approaches is used. The modified Lattice Boltzmann Model is used in [1].Here velocity if crowd is determined considering first few frames of the video. If there is sudden change in the velocity. Then it depicts the abnormal activity. It can be both used for global and local anomaly detection.

In [2] Optical flow algorithm is fused with deep auto encoder. Here normal video is converted into optical flow image. This is given as input to the deep auto encoder. Auto encoder is the encoder which converts the input pixels into 3 vector points which are plotted on the three-dimensional space. The normal behavior is extracted using auto encoder and are plotted on the three-dimensional space. If the point is outside this region then the situation is abnormal.

In [4] the classification technique is used is the two-class classification. Here the system is trained with both abnormal and normal events. It uses double sparse representation, which uses two dictionaries to store the normal events and abnormal events. The length of both dictionaries is fixed and is equal. The samples are classified into normal and abnormal during testing process. The uncertain samples are classified in the testing process. In [5], the main theme is that abnormal event changes non-uniformly compared to the normal event. The fixed temporal window size is chosen.The uniformity is checked using Pixel pair relationship, frame split. In [6] Scale Invariant Feature Transform (SIFT), weighted histogram and Hidden Markov Model are used. SIFT is better than optical flow method. It can be improved dividing the image into blocks to capture small local motions and involve more motion characteristics such as velocity and acceleration. In [7] Spatio-temporal feature and nonnegative locality constrained linear coding (NLLC) is used to detect abnormal events in surveillance video. The normal and abnormal events are classified by the Support Vector Machine. This method can perform well in both global abnormal and local abnormal event detection. In [8] Causal sliding window is used. When anomalies have tendency to change very fast then this method can be used. It is similar to any other windowing technique. Square and rectangular window are used. In [9] acceleration feature is used. Since it is independent of human detection and segmentation so it is robust. If motion fields are out of bounds then that mean the event is abnormal. The foreground is extracted by k method, detects the changes in foreground pixels while background is kept constant. But, the algorithm fails at high density crowd detection. In [10] this method the trained model consists of big data sheets with distributed storage environment for improving the performance of the intrusion Detection System. But, various distance computation errors may occur in this paper.

In [11] the process is dived into two parts. As for examine detection and suspicious behavior detection. Earlier morphological processing was used. The histogram of oriented gradients (HoG) feature with two-layer classifier to detect head-shoulder part for examinee detection. Sparse combination algorithm is employed for suspicious behavior detection for its fast-computing advantage. In [12], analysis of human behavior for people standing in queue is based on tracking first detection and tracking moving individual from stereo camera map real time. Activity zones are then automatically learn employing a soft computing based algorithms. [12] overcome the difficulty to perform with high human density and typical occasional motion by employed Bayesian filter. Errors due to occlusion or change of state of object isdrawback of this paper.

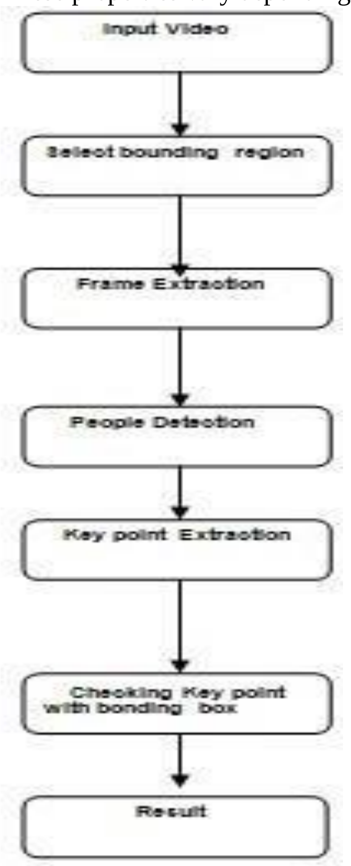
[13] uses a set of effectively computed easily interpretable scene level classification. The low dimensional descriptorvcombines two features from the crowd collectives and crowd conflict. Normal training data is available they use Gaussian Mixture to Model (GMM) for outline Detection Support Vector Machine (SVM) for binary classification. [13] usesoptical flow methods analyzing track-lets information to produce a low dimensional descriptor of scene level. [13] has drawback, which is crowd scale issues present cannot be addressed and no descriptor to label specific crowd behavior concepts. In [14] the requirements of public places crowd Stampede war warnings simulation systems were analyzed. Dense crowd has certain character subliminal increasing crowd density slow flow rate. This is detected in [14]. Aims at minimizing the stampede by locating the position of accidents and scatter the crowd through staff guidance basically early warnings. Uses of optical flow for analyzing parameters moving object identification methods mainly include background subtracting frames difference method single Gauss Background Modeling method. [14] is not accurate in huge crowd. [15] proposes a real-time detection using conditional random fields. It purposes bag of words (BoWs) to describe the motion information as the observation then apply CR and adaptive threshold to identify abnormal behavior. [15] claims CRF is best method to detect the anomaly. [16] This paper presents the 2D real human gesture grading state from monocular images based on OpenPose. This used motion trajectory to compute the equation which helps to judge the gesture.

III. Proposed System and Methodology

In every frame in the video, we detect keypoints of every person in the frame. At the time of installation of camera we create a bounding box. We continuously monitor the position of the keypoints of the person in the frame. When the keypoints move away from or into the bounding box we alert the system. As discussed in the overview keypoints are given by openpose. So, in our approach keypoints become the features of the system. These features are considered for the classification. The classification of the video is nothing but the movement of the keypoints which may go outside the frame which is being considered.

We have create a webcam object and then use the snapshot function to acquire images from the webcam. We can acquire a single image or set up a loop to acquire multiple images. Use the preview function to test the camera stream.

We can set device-specific properties, such as resolution and brightness, if our camera allows programmatic access. These properties vary depending on our device.



The initial studies in abnormal event detection focused on trajectory analysis, where a moving object is considered as abnormal if its trajectory doesn't respect the fitted model during the training period. The main limits of such method are the sensitivity to occlusion and the effectiveness of detecting abnormal shapes with normal trajectories. These methods can be used in detecting deviant trajectories in scenes containing few objects but not achieve satisfactory performance for other applications.

A. Face Detection using Viola-Jones method

The Viola-Jones algorithm is a widely used mechanism for object detection. The main property of this algorithm is that training is slow, but detection is fast. This algorithm uses Haarbasis feature filters, so it does not use multiplications.

The efficiency of the Viola-Jones algorithm can be significantly increased by first generating the integral image.

integral image.

$$II(y,x) = \sum_{p=0}^{y} \sum_{q=0}^{x} Y(p,q)$$

The integral image allows integrals for the Haar extractors to be calculated by adding only four numbers. For example, the image integral of area ABCD (Fig.1) is calculated as $II(y_A,x_A) - II(y_B,x_B) - II(y_C,x_C) + II(y_D,x_D)$.

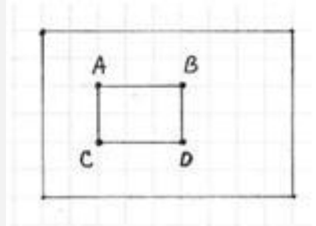


Fig.1 Image area integration using integral image

Detection happens inside a detection window. A minimum and maximum window size is chosen, and for each size a sliding step size is chosen. Then the detection window is moved across the image as follows:

1. Set the minimum window size, and sliding step corresponding to that size.
2. For the chosen window size, slide the window vertically and horizontally with the same step. At each step, a set of *N* face recognition filters is applied. If one filter gives a positive answer, the face is detected in the current widow.
3. If the size of the window is the maximum size stop the procedure. Otherwise increase the size of the window and corresponding sliding step to the next chosen size and go to the step 2.

Each face recognition filter (from the set of *N* filters) contains a set of cascade-connected classifiers. Each classifier looks at a rectangular subset of the detection window and determines if it looks like a face. If it does, the next classifier is applied. If all classifiers give a positive answer, the filter gives a positive answer and the face is recognized. Otherwise the next filter in the set of *N* filters is run.

Each classifier is composed of Haar feature extractors (weak classifiers). Each Haar feature is the weighted sum of 2-D integrals of small rectangular areas attached to each other. The weights may take values ±1. Fig.2 shows examples of Haar features relative to the enclosing detection window. Gray areas have a positive weight and white areas have a negative weight. Haar feature extractors are scaled with respect to the detection window size.

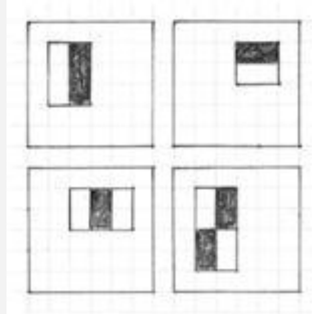


Fig.2 Example rectangle features shown relative to the enclosing detection window

The classifier decision is defined as:

$$C_m = \begin{cases} 1, & \sum_{i=0}^{l_m-1} F_{m,i} > \theta_m \\ 0, & \text{otherwise} \end{cases}$$

$$F_{m,i} = \begin{cases} \alpha_{m,i}, & \text{if } f_{m,i} > t_{m,i} \\ \beta_{m,i}, & \text{otherwise} \end{cases}$$

$f_{m,i}$ is the weighted sum of the 2-D integrals. is the decision threshold for the *i*-th feature extractor. $\alpha_{m,i}$ and $\beta_{m,i}$ are constant values associated with the *i*-th feature extractor. θ_m is the decision threshold for the *m*-th classifier.

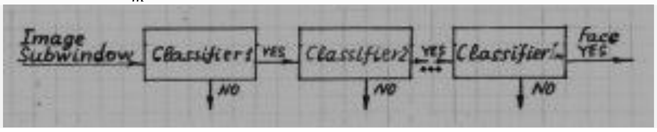


Fig.3 Object detection Viola-Jones filter

The cascade architecture is very efficient because the classifiers with the fewest features are placed at the beginning of the cascade, minimizing the total required computation. The most popular algorithm for features training is AdaBoost.

Automatic human detection and tracking is an important feature of video surveillance systems. It can improve a system's performance in fields such as security, safety, human activity monitoring etc. Human detection systems can have different goals such as detecting the presence of humans, recognition of abnormal behavior (falls, climbing, running, etc.), identification of specific individuals, etc.

B. Human Detection and Tracking in Video Surveillance System

For human detection and tracking, the surveillance system should represent the video as a sequence of still images.

Frame subsampling could be done. For this sequence, the background has to be modeled and then subtracted, and the foreground extracted. Reduction of noise and image compression artifacts can be very useful. Then, human detection may be done based on shape detection and motion from previous frames. From this information, regions of interest (ROI) can be identified.

One popular technique for human shape classification is histogram of oriented gradients (HOG). Once the ROI is classified and human detection is confirmed, human tracking can start. Human tracking has to be able to handle occlusion events. Finally, human activity analysis can happen. For that, human body parts must be classified and labeled. Based on human body part motion, the human activity analysis can happen. If abnormal behavior is detected, an alarm can be triggered. To identify individuals, face detection and recognition can also be applied.

C. Pseudo Code

This section explains the algorithm that is being used in this approach of anomaly detection.

```
Start
WHILE video is running Take frames continuously
IF TRUE select first frame
Select Region of Interest
ENDIF
Generate keypoints
IF keypoints outside the ROI
Detect as anomaly
ELSE
Normal event
ENDIF
ENDWHILE
```

IV. Results

Fig.4:Identification of image from video

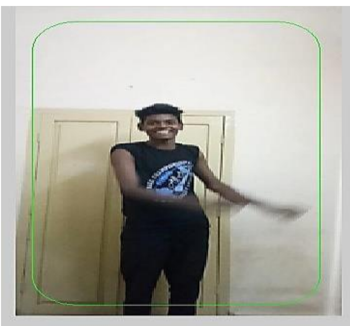


Fig.5 Selection of boundary region

Fig 6: Face detection

Fig 7: Body detection

Fig 8:Key points Recognition

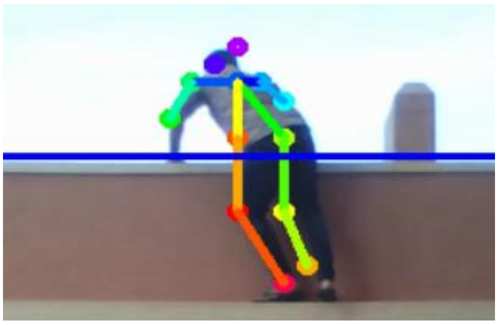


Fig 9: Anomaly detection

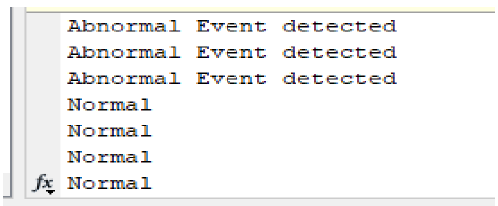


Fig 10:Ouput(Alert message)

V. CONCLUSION AND FUTURE SCOPE

The anomaly detection has lot of application. This paper was able to detect an anomaly in the video using novel approach of key keypoint detection. The proposed approach worked efficiently on new generation laptop with medium to high end processor. However, the performance was not great with the laptops with older Processor.

In future work, the keypoint detection can be not only for pose detection but also for human tracking, as tracking with the key points will be very efficient compared to traditional methods. The present system cannot take decisions on its own. It requires high performance laptops for the process to run smoothly.

Future Enhancements:

- Training system with artificial intelligence to make decision of it's own.
- Implementation of this system as a cloud based application, just by feeding the video stream directly through web based service,which is possible to detect anomalies.

ACKNOWLEDGMENT

I would like to express my regards and acknowledgement for all those who helped us in carrying out this paper in perfect manner .I am indebted to Mr.A.Thyagarajamurthy,Prof in Department of Electronics and Communication, JSS Science and Technology University, for valuable suggestions and constant encouragement provided for successful completion of thispaper.

REFERENCES

[1] Yiran Xue, ,Peng Liu, Ye Tao, Xianglong Tang, "Abnormal Prediction Of Dense Crowd Videos By A Purpose Driven Lattice Boltzmann Model",Int. J Math Comput. Sci. 2017 , Vol. 27, No. 1, pp. 181-194, February2017

[2] Meina Qiao, Tian Wang, Jiakun Li1, Ce Li, Zhiwei Lin, Hichem Snoussi,

[3] "Abnormal Event Detection based on Auto encoder fusing opticalflow ,Proceedings of the 36th Chinese Control Conference", July2017.

[4] Dinesh Singh, C. Krishna Mohan, "Graph formulation of video activities for abnormal activity recognition", Visual Learning and Intelligence Group (VIGIL), ScienceDirect, Pattern recognition, vol. 65, pp. 265-272, January 2017

[5] Peng Liu, Ye Tao, Wei Zhao, Xianglong Tang, "Abnormal crowd motion detection using double sparse representation",Pattern Recognition Research enter,Neurocomputing, Vol. 269, pp. 3-12, June2017

[6] Kaelon Lloyd, Paul L. Rosin, David Marshall, Simon C. Moore, "Detecting violent and abnormal crowd activity using temporal analysis of Grey level co-occurrence matrix (GLCM)-based texture measures", Machine Vision and Applications, Vol. 28. ,pp. 361-371, March2017

[7] Dongping Zhang, Kaihang Xu, Huailiang Peng, Ye Shen. "Abnormal Crowd Motion Behaviour Detection based on SIFT Flow", International Journal of signal processing and pattern recognition", Vol. 9, No. 1, pp. 289- 302, September2016

[8] Yu Zhao, Lei Zhou, Keren Fu, Jie Yang, "Abnormal Event Detection Using Spatio-Temporal Feature And NonnegativeLocality-Constrained Linear Coding", IEEE Conference, Vol. 9, pp. 3354-3358, June 2016

[9] Chein-I Chang, Yulei Wang,and Shih-Yu Chen, "Anomaly Detection Using Causal Sliding Windows", IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, Vol. 7, pp. 3260-3270, July2015

[10] Chunyu Chen, Yu Shao, and Xiaojun Bi,"Detection of Anomalous Crowd BehaviorBasedontheAccelerationFeature",IEEESensorsJournal,Vol.15, No. 12, pp. 7252-7261, December 2015

[11] jabez Ja, Dr.B.Muthukumar, "Intrusion Detection System (IDS): Anomaly Detection using Outlier Detection Approach", International Conference on Intelligent Computing, Communication and Convergence(ICCC-2015), ScienceDirect, Vol. 48, pp. 338-346, April2015

[12] Miaomiao Ding, Jiahui Zhao, Fangyu Hu, "Abnormal Behavior Analysis Based on Examination Surveillance Video", 9th International Symposium on Computational Intelligence and Design, Vol 16, pp. 2473-3547, June2016

[13] Luis Patino and James Ferryman University of Reading, "Abnormal behaviour detection on queue analysis from stereo cameras", Computational Vision Group White knights, Vol 3, pp. 345-350, July2015

[14] Mark Marsden Kevin McGuinness Suzanne Little Noel E. OConnor, "Holistic features for real-time crowd behaviour anomaly detection", IEEE conference, Vol. 9 , pp. 918-922, June2016

[15] Shangnan Liu, Qiang Cheng, Zhenjiang Zhu, Hao Zhang, "Analysis and Design of Public Places Crowd Stampede Early-Warning Simulating System",International Conference on industrial Informatics, Vol 978-1-5990- 3575-5, pp. 210-213, May2016

[16] Ben-Syuan Huang, Shih-Chung Hsu and Chung-Lin Huang, "Abnormal behavior detection using conditional random fields",Vol 8, pp.235-240, January 2016

[17] Real-time human gesture grading based on OpenPose, Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2017 10th International Congress, 14-16 Oct. 2017, 10.1109/CISP- BMEI.2017.8301910