Literature Review

## Multiple Anomalous Activity Detection in Videos

[2017] Sarita Chaudharya, Mohd Aamir Khana, Charul Bhatnagara

In the paper the authors use a Rule-Based approach to classify events as normal and anomalous. Motion patterns are analyzed and used as features. They followed the concept of dominant set in which dominant behavior (frequently occurring) is treated as normal and less dominant behavior is considered as anomalous.

The Rule-Based method is a semi-supervised method, so it only requires normal data for training the system. In Rule-based method, rules are established, and those rules help to categorize the samples as normal or anomalous. Samples that match the rules are treated as normal and those not matching the rules are treated as abnormal or anomalous samples.

They considered running, walking and crawling as main activities of their approach. In which behavior pattern and direction of object help to identify anomalous events. The proposed framework consists of the following components: determining object direction and analyzing motion pattern.

The main components of the framework are: preprocessing phase, feature extraction phase and recognition phase. In first phase, the preprocessing, a moving object is detected and noise removal is done. In the feature extraction phase different features like centroid, movement, speed, direction and dimensions/measurements are calculated. Finally, in last phase Rule-based classification method is used to classify the activities of the input video and alarm is generated for every suspicious activity.

Gaussian Mixture Model (GMM) is used for object detection (background subtraction) and some problem domain rules are used to distinguish different types of behavior.

This method doesn’t require any type of labeled or unlabeled training samples. They require some external knowledge or rules of the domain to create the model. The accuracy of this approach is heavily dependent on rules. In this method set of rules, pre-defined threshold is defined initially. It includes if-then rules for making decisions and these rules decide whether the recognized activity belongs to the normal activity class or anomalous activity class.

Dataset – due to unavailability of standard dataset the authors created a new dataset. The dataset consists of 45 videos of three activities: walking, running and crawling of multiple people.

Tests were performed to evaluate the different types of activities.

## A fast recognition algorithm for suspicious behavior in high definition videos

[2015] Chundi Mu, Jianbin Xie, Wei Yan, Tong Liu, Peiqin Li

In this method, the motion vectors in the video streaming are taken into account. The motion vectors are extracted from the video streaming directly. Then, the moving target region is extracted by optimal threshold method. The direction and velocity parameters are then extracted from the motion vectors’ modulus. A support vector machine (SVM) is used to learn and classify the input videos.

A moving object segmentation algorithm is proposed based on the modulus of motion vectors. This method uses velocity to distinguish interesting targets from interference targets (e.g., ambient noise). Seven features are proposed to describe each moving target, which makes it easier to judge the state of the target.

Through learning and analyzing the vectors from the data stream, the proposed method can identify suspicious behavior such as wandering, trailing, chasing, and falling down.

All of the vector features are extracted from regional motion vectors.

This paper uses an optimized adaptive threshold segmentation algorithm, which has strong anti-noise performance, high speed, and high efficiency. This algorithm effectively reduces the false alarm rate.

There is no need to construct a background model for this algorithm.

If there exists a single target suspicious behavior (e.g., wandering and falling down), the direction and trajectory of motion vectors show certain regularity. When there exists a wandering suspicious behavior in the video, the motion vectors of the target are periodically changing. For a man falling down, the speed and direction of the target would change. For suspicious behavior of multiple targets, such as following and chasing, direction and velocity have different characteristics. For following, the direction and velocity of targets are always the same. For chasing, the direction of the targets is the same, and the targets always have a great velocity. Therefore, the authors focused on analyzing the direction and velocity of the target. To analyze the features of the target more precisely, a new concept called motion vectors in region (MVR) is defined. Each MVR corresponds to a moving region.

7-D features {­θ, V, σθ, σV, E­θ, EV, InterDj} can be employed to describe the targets extracted from the frame. θ and V represent the average direction and velocity of the target; σθ and σV represent the direction variance and the velocity variance; E­θ and EV represent entropy of direction and velocity; InterDj interesting degree of inter-frame InterD. For these suspicious behaviors was detected in the paper, that the direction, velocity and interframe difference are the most important features of them.

SVM is used as the classifier. Gaussian radial basis function is used.

The authors built a suspicious behavior database for this paper in particular. This database consists of five human action classes: wandering, trailing, chasing, falling down, and normal activity. Total of 300 HD video samples.

For the first experiment, videos were divided into training samples and testing samples with both normal and suspicious behavior. One-to-one classifiers were generated, they can only classify input videos as suspicious or not. However, the type of suspicious behavior must be identified. Therefore, a second experiment is conducted, which uses a voting method to determine the category of the suspicious behavior. In this experiment, 100 samples for training SVM classifiers (every category has 20 samples) are selected. Ten classifiers are obtained. When a new video is input, it will be classified by these ten classifiers separately, and the one with the most votes is selected as the result.

## Expert video-surveillance system for real-time detection of suspicious behaviors in shopping malls

[2015] Roberto Arroyo, J. Javier Yebes, Luis M. Bergasa, Iván G. Daza, Javier Almazán

This paper focuses in the detection of suspicious behavior in shopping malls. Situations that must be analyzed include a store entry and exit, loitering events that can end in theft and situations where a cash desk in unattended.

The paper presents a comparison between the main background subtraction methods and concludes that Gaussian mixture model by Zivkov (2004) is the best for this specific application. After implementing the method the segmented images are filtered by applying dilation and erosion.

The foreground objects, distinguished as blobs, are filtered by size and positional factors and grouped by their blob fusion algorithm. The blob fusion technique proposed consists of combining foreground blobs identified in a frame which can potentially correspond to one person

An innovative tracking method is proposed, while other methods are mainly based on trajectories, this method also uses visual appearance information in occlusion situations for improving the performance of the method. This method consists in a two-step algorithm:

1. The detected objects to track are matched along the video sequence employing an optimization approach based on solving the association problem as a LSAP (Linear Sum Assignment Problem) and considering the estimations of a Kalman filter (to predict the future positions and sizes of the tracked objects based on a constant velocity model);

2. Occlusions between objects are managed applying a method based on visual appearance, in which several image descriptors are tested: GCH (Global Color Histogram) – low computational cost descriptor based on the color histogram of an object, differentiates people by the colors of their clothes; LBP (Local Binary Pattern) – powerful texture descriptor also employed in human appearance with an efficient performance; and HOG (Histogram of Oriented Gradients) – describes the shape of a person through its gradient distribution, having a more expensive computational cost. To compare the appearance features of a person before and after of an occlusion, SVM kernel functions are implemented.

The high-level view of the proposed system consists of a multi-camera approach where three differentiated shop zones are under monitoring. Each zone has a particular situation to evaluate: entry or exit of people in the entrance zone (analysis of trajectory), loitering events in the shop interior (risk zones are defined and loitering is defined as the presence of an individual in an area for a period of time longer than a given time threshold) and unattended cash desk situations in the payment area.

The experiments carried out are divided into two groups: the test of the proposed tracking algorithm and the analysis of the methods proposed for alarm detection in shops. For tracking results evaluation, the public CAVIAR dataset is employed. For the test of the alarm detection in shopping malls a private dataset is used, which is composed of videos with naturalistic shop situations and conditions as low illuminated zones, shadow effects, noise, frame loss and crowds. The tracking experiment shows superior results when compared with other algorithms.

## Local abnormal behavior detection based on optical flow and spatio-temporal gradient

[2015] Songhao Zhu, Juanjuan Hu, Zhe Shi

In the proposed method, firstly, a video sequence is divided into spatio-temporal blobs; then, a statistical method based on the semiparametric model is adopted to detect these blobs where abnormal behaviors are most likely to appear; finally, maximum optical flow energy and local nearest descriptor are utilized to determine whether these suspicious blobs really contain abnormal behaviors.

In the method, first, bag of words is adopted to describe the optical flow information of each blob, and the semi-parametric based model is adopted to detect suspicious local abnormal blobs; then, suspicious abnormal blobs are divided into rectangular cells equally, and the maximum optical flow energy method is adopted to detect undoubted abnormal cells; finally, the local nearest descriptor and Mixed Naïve Bayes model are adopted to determinate anomaly behaviors.

Each video image frame is first divided into blobs. A blob is considered a suspicious abnormal blob according to a likelihood probability model.

Compared with blobs images of normal behavior, blobs images of abnormal behavior often have obvious characteristics of higher motion velocity and more disordered motion direction. Therefore, the optical flow energy of an abnormal behavior blob is larger than that of a normal behavior blob. The basic idea of local abnormal behavior detection in this paper is described as follows: the mixed naive Bayes model is first utilized to train nearest neighbor descriptor of normal cells, and then the trained nearest neighbor descriptor is utilized to determine whether each test cell is an abnormal cell or not.

Two datasets are used for testing. The UCSD dataset includes ped1 and ped2 subsets for detecting local abnormal behavior. These subsets have training, validation and testing video sequences. The normal behaviors on the UCSD datasets are defined as walking with normal speed. The local abnormal behaviors mainly include irregular moving, such as skating, biking, and driving.

The subway dataset contains two video sequences recorded by a camera at the entrance gate and a camera at the exit gate respectively. The first video sequence contains normal behaviors including going down through the turnstiles and entering the platform. Abnormal behaviors include walking in the wrong direction, irregular interactions between people, sudden stopping, running fast. The second one, the exit gate surveillance video, contains the anomalous events of walking in the wrong direction and loitering near the exit gate. These videos are not labeled as training or testing videos.

The proposed method showed to have better results than other methods.