

# A OUTLINE OF SUSPICIOUS AND VIOLENT ACTIVITY DETECTION OF HUMANS

Karthiga.D, Dr.M.Arumaiselvam

Research Scholar, HoD

PG & Research Department of Computer Science,

ST.Joseph's College of Arts & Science (Autonomous), Cuddalore, India.

**Abstract:** Crimes like stealing, violence against folks, injury to property, etc., became quite common in a very society, that a significant concern. The normal police work systems act like post mortem tools within the sense that they'll be used for the investigation to detect the person behind the theft, but it is only after the crime has already occurred. During this chapter, we tend to propose a technique for mechanically detection the suspicious or violent activities of an individual from the police work video. We tend to train the SVM classifier with the HOG options extracted from the video frames of two types: frames showing no violent activities and people showing violent activities like kicking, pushing, punching, etc. In the testing phase, the frames from the surveillance video square measure browse and processed so as to classify them as violent or traditional frames. If the frames classified as violent frames square measure detected, associate alarm is raised to alert the controller. It will be wont to keep track of the time period that an individual is found loitering at a place being monitored. If the time exceeds a predefined threshold, the alarm is raised to alert regarding any potential suspicious activity in order that it will be checked on time.

**Keywords:** police investigation, video tracking, violence activity, svm, hog

## I.Introduction

Nowadays, the crime and violence have inflated in society in a very massive manner, requiring someone to hunt safety and security at every stage of his life. Police investigation camera systems square measure accessible to stay a watch on the live happenings in a very region of interest. Observation such places exploitation cameras square measure less expensive than using human resource World Health Organization will unceasingly monitor the live footage for any violent or suspicious action. Even though one manages to rent person(s) to try such task, the efficiency of someone in doing it might be at risk of errors as 82% of security observers have multiple co-occurring duties like checking-in guests, attending calls, etc. Thus, there's a desire for associate intelligent autonomous closed-circuit television which will discover the suspicious and violent activities of humans in police investigation videos to forestall injury on time.

In this chapter, associate acceptable algorithmic program is projected to satisfy the higher than would like. Though such systems are developed, however it's difficult as a result of there square measure several problems that require be self-addressed. One in every of the problems to be self-addressed is however intelligent is that the system to grasp the human behaviour in a region being monitored. Our work focuses on higher understanding of human-movement preponderantly suspicious behaviour, so as to acknowledge associate persona non grata exploitation the human movement detection algorithmic program. Also, the actions like kicking, pushing, punching, etc., square measure treated as violent/abnormal activities associated an alarm is raised whenever such acts of violence square measure in method so no any injury will present itself. We analyze, discover and track the motion of someone in a very given police investigation video and based mostly upon it we tend to reason whether or not there square measure some sorts of acts of violence, viz. Kicking, pushing, punching or a potential suspicious activity like loitering or looking around with malicious intent. To discover acts of violence, a support vector machine (SVM) classifier is pre trained with many samples of acts of kicking, pushing and punching. Once the classifier is trained, each frame of the video is provided to the classifier to discover if there's any violence of the above-named classes. If the classifier returns positive result classifying the frame mutually within which some act of violence is found, then

associate alarm is raised to alert the observer. We will use it to cypher the time that specific a selected person is gift in a very particular frame loitering there. If the time therefore calculated for the person exceeds a pre-determined threshold, then associate alarm is raised indicating a possible suspicious activity. Thus, the period of time detection of suspicious human movement will be wont to overcome the shortcomings of the prevailing systems and to produce a correct automatic closed-circuit television.

## II.METHODOLOGY

The basic steps of a typical video closed-circuit television square measure as follows

- Building a background model
- Foreground element extraction
- Object segmentation
- Object classification
- Object trailing
- Action recognition

The terribly step is building a background model whose sole purpose is to represent the atmosphere within the absence of foreground objects. In literature, there are several strategies for building the background models [2]. That differ within the ability to induce updated to reflect the alterations within the atmosphere. These strategious area unit generally as adjustive and non-adaptive techniques. The foreground pixel extraction step separates the pixels of a picture that aren't a part of the background model.

This difference forms the premise for any analysis in consequent steps. A number of the strategies for foreground pixel extraction are background subtraction, optical flow, temporal differencing, etc. [4–6]. Object segmentation step teams the similar foreground pixels into same regions victimization some similarity metrics. A number of the similarity metrics for object segmentation ar location primarily based, color primarily based, proximity primarily based or hybrid of some or all of the higher than [7]. The fourth step is object classification. There may be many styles of moving objects within the space being monitored and a different pursuit technique is also used for every sort. As an example, the moving objects is also humans, vehicles or objects of interest of Associate in Nursing investigation application. Thus, the item the item be mentioned as normal pattern recognition task. Two main classes of moving object classes are motion based classes [10] and form based classes [7]. Different form info of motion regions like boxes, points, blobs and silhouettes are used for the higher than higher than task. Future step is object pursuit that may as a way of locating the object(s) in motion over a time period, primarily matching or associating the target objects in consecutive frames. Low video frame rate as compared to the speed of object motion poses an excellent challenge in object pursuit. Matters becomes worse once the background has objects similar in options thereto of the foreground objects, or there are objects that cause severe occlusion to the item being half-tracked. The most classes of object pursuit strategies embrace region primarily based, contour primarily based, feature primarily based, model primarily based and their hybrid forms. The final step is action recognition, that is roughly a method of distinguishing the actions so as to know what is occurring in the environment[10]. In an automatic video surveillance system, there is an excellent sensible importance of suspicious human behaviour [3]. The popularity is achieved through a mixture of image process and process intelligence approaches. The previous is employed to extract low level image options whereas the later provides knowledgeable choices. There exist many works on low level image options like object detection, recognition and pursuit, however few give reliable classes and analysis of human activities from the video frame sequences. The most steps in our work may be listed as follows:

1. To extract frames from coaching police investigation video.
2. To arrange ground truth.
3. To extract HOG (histogram of orientated gradients) options from frames.
4. To come up with coaching feature matrix.
5. To arrange category label matrix for coaching feature matrix half-dozen.
6. To coach SVM (support vector machine) classifier.
7. To sight motion in check (input) video.

8. To trace and predict the movement of every person in video.
9. To send each frame of input video to SVM classifier.
10. To boost alarm if check frame is classified as frame with violent activity.

We now elaborate these steps.

#### 2.1 To extract frames from coaching police investigation video:

We extract the frames of the coaching police work video using Video Reader perform in matlab and store them in .jpg format.

#### 2.2 To arrange ground truth:

The frames therefore extracted from the video square measure analyzed and classified into frames containing traditional activity or any variety of violent activity, viz. Kicking, pushing, punching, etc.

#### 2.3 To extract HOG (histogram of orientated gradients) options from frames:

The bar graph of familiarized gradient could be a feature descriptor accustomed objects and that they square measure principally fitted to detection of humans in pictures. The HOG descriptor technique counts the incidence of gradient orientation in localized parts of the image—detection window, or the region of interest.

- The image is split into tiny connected regions, mentioned as cells. For every cell, a bar graph of gradient direction for the pixels is calculated.
- Each cell is discretized into angular bins in accordance to the gradient orientations. There's a contribution of the weighted gradient to the corresponding angular bins for every cell.
- Blocks square measure thought-about as spatial regions of teams of adjacent cells. Formation of the blocks is the basis for grouping and normalizing the histograms.
- Block bar graph is pictured by the normalized cluster of histograms. The descriptor is successively pictured by the set of those block histograms.

#### 2.4 To come up with coaching feature matrix:

The HOG options are generated for the traditional frames still because frames showing violent activity to get the feature matrix. This feature matrix is employed for coaching the SVM classifier. The HOG options generated for every frame is of size  $1 \times 189036$ .

#### 2.5 To arrange category label matrix for coaching feature matrix half-dozen:

A class label matrix is created for each kind of frame. The class label matrix is used for training the SVM classifier.

#### 2.6 To coach SVM (support vector machine) classifier:

The SVM classifier was developed by Vapnik (A training algorithm for optimal margin classifier). Victimisation the coaching feature matrix and also the category label matrix, we tend to train the SVM classifier that predicts the category label of associate degree input take a look at video frame. Classification is one in all the sub-parts of machine learning algorithms. The support vector machine could be a supervised learning model that's used for knowledge knowledge and regression analysis. Consider samples from two classes that are mapped in space. The SVM tries to find a linear call surface (hyperplane) that may separate the samples of those categories in such some way that the border line samples of every category have the most important distance or gap or margin between them. These border line samples of every category area unit referred to as the support vectors. Just in case such linear call surface doesn't exist, the information is mapped into a way higher dimensional feature house, wherever the separating hyperplane is found. The feature house is also made victimisation the mathematical projection, one in all them is kerneltrick.

- It is strong to sizable amount of variables and comparatively smaller number of samples.
- It will learn easy to extremely advanced information models.

- It will create use of subtle principles of arithmetic to avoid over fitting.
- It produces superior empirical results.

## 2.7 To sight motion in check (input) video:

We input to take a look at video to the most module that reads the video frame by frame and also the motion of the person(s) in video is detected by background subtraction mistreatment the reconciling mathematician mixture model. The binary image therefore obtained is usually rip-roaring. The noise is removed mistreatment the morphological operations like erosion and dilation. Finally, the blob analysis detects the teams of connected pixels, that are probably to correspond to moving objects.

**Background Subtraction:** it's a method to extract the foreground of a picture for more analysis. Straight forward background subtraction techniques like, mean and median filtering area unit quick however their international and constant thresholds build them build for the difficult real-world issues. The adjustive background mixture model handles the difficult things like bimodal backgrounds, long-run scene changes and repetitive motions in litter. It will more be improved by incorporating the temporal data, or victimization some regional background subtraction in conjunction with it.

**Morphological Operations:** The dilation operator takes two items of information as input. The first is the image that is to be expanded and also the second could be a little set of coordinate points, referred to as a structuring part. The essential dilation operator on a binary image is to bit by bit enlarge the boundaries of regions of foreground pixels (generally white pixels). Thus, the areas of foreground pixels grow in size whereas the holes among those regions become smaller. So as to reason the dilation of a binary image by a structuring part, we tend to take every of the background constituents within the input image one by one and put the structuring part on the binary image in such how that the origin of the structuring part coincides with the input pixel position. If a minimum of one pixel component constituent element within the structuring element coincides with a foreground constituent within the image, then the input constituent is about to the foreground price. If all the corresponding pixels within the image area unit background, the input constituent is assumed because the background price. The erosion operator takes two items of information as input. The first is the image that is to be scoured and also the second could be a little set of coordinate points, referred to as a structuring part. The essential the essential the erosion operator on a binary image is to erode the boundaries of regions of foreground pixels (generally white pixels). Thus, the areas of foreground pixels shrink in size and also the holes among those areas become larger.

**Blob analysis:** This method is employed to isolate the blob(s) (objects) in an exceedingly binary image, calculate its varied options like space, centroid, bounding box coordinates, etc., then it uses them to classify the blobs as per the wants. A blob consists of a gaggle of connected pixels. To find out whether the pixels square measure connected, we have tendency to observe their property that tells that pixels square measure neighbors and that aren't. The two most typical forms of property are: 4-connectivity and 8-connectivity. The 8connectivity is a correct than the 4-connectivity, however the 4-connectivity is usually applied since it needs fewer computations and thus will method a picture quicker. There square measure variety of algorithms for finding the blobs and these algorithms square measure usually known as connected part analysis or connected part labelling. Example of such an algorithm is the Grass-fire algorithm. Space of a blob is that the variety of pixels lined by the blob and therefore the space feature is employed to get rid of blobs that square measure too huge or too tiny from the image. Bounding box of a blob is that the minimum parallelogram that contains the blob. It's defined by researching all pixels for a blob and finding the four pixels with the minimum x-value, most x-value, minimum y value and most y-value, severally. A bounding box is used as a ROI (region of interest). The bounding box magnitude relation of a blob is defined because the height of the bounding box divided by its dimension. Compactness of a blob is defined because the magnitude relation of the blob's space to the world of the bounding box. This can be accustomed distinguish compact blobs from non-compact ones. In our work, the blob analysis is employed within the method of detection of the moving persons in an exceedingly police investigation video.

## 2.8 To trace and predict the movement of every person in video:

We use the Kalman filter with constant rate to trace and predict someone in an exceedingly video. When someone goes out of the frame, its track is deleted from memory. The short lived tracks square measure unnoticed the reliable tracks square measure forbidden. The Kalman filter is employed where we've got unsure data regarding some dynamic system and that we will create associate degree intelligent guess regarding what the system goes to try and do next. Although there's interference with the clean motion we have a tendency to guessed regarding, the Kalman filter is nice is nice out what truly happened. It will profit of the correlation between phenomena that one wouldn't have typically thought of exploiting. The Kalman filters perform well for dynamic systems. They're light-weight on memory as they are doing not have to be compelled to keep any history aside from the previous state, and therefore square measure in no time, creating them like minded for real time issues and embedded systems.

## 2.9 To send each frame of input video to SVM classifier:

While reading the input video frame by frame, we have a tendency to contemplate each frame of the check video for prediction by the trained SVM classifier model so the frame are often classified as either traditional frame or one that has violent activity in it associated consequently an alarm are often raised to alert the human observer at the opposite finish. We have a tendency to send every fifth frame for prediction instead of all the frames of the test video because checking all the frames would build the detection system lag as prediction by SVM model will take your time for outputting the expected label for a check frame.

## 2.10 To boost alarm if check frame is classified as frame with violent activity:

The SVM classifier predicts the label of a check frame and tells whether it's traditional frame or the one which shows some violent activity. If it is of the second type, then the system raises associate degree alarm within the type of a popup message box and a few sound to alert the human observer sitting within the room. There's another alert mechanism for this technique that keeps a track of however long a specific person seems within the region being monitored. If the closing date for someone exceeds a definite threshold worth (which is pre-determined in our case and might be set to any worth supported the appliance and region being monitored), then the system treats the person loitering associate degree alerts the human observer within the room regarding the potential suspicious behaviour of that person by raising an alarm within the type of a popup message box and a few sound. There are primarily two types of learning: supervised and unsupervised learning.

### Supervised Learning

- In this case, the patterns area unit discovered within the information that relate the information attributes with the target category attributes. It's like learning within the presence of associate degree skilled or teacher.
- Coaching information is tagged with a category or worth, which implies that the expert/teacher provides with correct category of the coaching information and also the system learns from it. Any error is reported as feedback to the system and also the system updates itself consequently. The coaching is continual over and over till the system achieves a desired expertise level to properly classify the check information.
- The most goal here is to predict a category or worth label of the check information. The supervised learning algorithms embrace call trees, support vector machine, neural networks, theorem etc.
- In our work, the system is trained with samples of the frames portrayal the violent activities then a model is developed that classifies a frame of a check video as traditional or violent one.

Thus, it's supervised learning within the sense that the system is first shown that one is traditional and that one is violent frame. As additional and additional style of coaching frames area unit on the market, the system develops the power to classify the frames additional accurately.

## Unsupervised Learning

- It is like learning within the absence of a coach. The system learns on its own by exploring the coaching knowledge to find some intrinsic structures in it. Thus, it's self guiding learning rule with internal self analysis against some criteria.
- There's no information of output category or price. The coaching knowledge is unlabeled or the worth is unknown.
- The most goal here is to see knowledge patterns or grouping. Samples of algorithms mistreatment unsupervised learning are: k-means, genetic rule, clump approaches.
- If unsupervised learning is employed in our suspicious violent activity detection from the police investigation videos, then the system forms two clusters by itself: one amongst the traditional frames and different one amongst frames portraying violent activities by observant an oversized range of coaching videos and supported the acceptable options to differentiate within the clusters shaped. Once take a look at frame is input for once, it's going to provides a probabilistic output indicating the chance of it happiness to every of the two clusters shaped, rather than a crisp output within the style of one cluster it belongs to using unsupervised learning, it becomes more efficient over time as it witnesses a lot of coaching videos. It is disadvantage is that it wants a great deal of coaching videos to present an accurate an accurate, increasing the coaching time.

## III. Experimental Results

The proposed system can detect the violent activities fairly well and the motion detection is additionally quite correct. The person moving is being enclosed by a parallelogram together with a trailing range displayed on the highest of the bounding box (rectangle). The corresponding background ablated video is shown together with the first one. We've got used a complete of 835 video frames from the UT-Interaction dataset, ICPR contest on linguistics Description of Human Activities (SDHA), for coaching the SVM classifier. The HOG feature vector generated for every frame is of size  $1 \times 189036$  and so the entire coaching feature matrix is of size  $835 \times 189036$  as there are 835 frames. For testing purpose, we've got input a video to the system and if there's any frame gift within the video showing any of the violent activities, the system pauses there to lift AN alarm so as to alert the observer concerning it. The observer decides what to try and do next whether or not to require action responding to the alarm raised or to let the police investigation proceed treating the alarm simply raised as a warning. Just in case a selected person has been half-tracked during a video for a period bigger than a pre-determined threshold, an alarm is raised voice communication the person is loitering which there could also be a possible suspicious activity.

## IV. Future Work

Although our system performs a tendency toll and achieves nearly 89% accuracy in detective work suspicious and violent activities of humans in police investigation videos; we but prefer to define a number of the longer term work that we'd undertake to boost and build the system additional sturdy.

We would prefer to use additional correct optical flow options along background subtraction for motion segmentation. The optical flow options square measure quite common for assessing motion from a collection of pictures, that is AN approximation of the two dimensional flow from the image intensities. It's the speed field, that wraps one image into another terribly similar image.

We have used police investigation videos recorded from the only camera for the system. We'd we would system to produce additional correct and sturdy results by victimization videos of constant region of interest captured by multiple cameras unbroken at unbroken angles. For this, we'd like to possess a synchronization and standardisation module which might facilitate these cameras add coordination to provide additional correct detection of suspicious and violent human activities.

The feature matrix that we've generated to coach our SVM classifier is sort of giant with its dimension being  $835 \times 189036$  and also the memory occupied by it on disc is around 650mb. So, we would like to use a feature reduction algorithm to minimize the quantity of relevant options therefore on reach twin advantage viz. Smaller feature matrix with no compromise on the coaching quality. With smaller feature matrix, the

system will run a lot of quicker with improved accuracy. Aside from HOG options, we'd additionally prefer to use alternative options like form descriptors in order that the human activity is detected more accurately and the number of false alarms is greatly reduced. We'd prefer to train a cascade of different forms of detectors/classifiers instead of simply victimization the SVM classifier in order that the prediction of the class of a frame could be a ton additional correct than within the current system.

Currently, we've simply unbroken track on the number of your time an individual is found moving incessantly solely within the region being monitored by the police investigation camera. If the time exceeds a pre-determined threshold, then the system flags the person as loitering, indicating some potential suspicious act is also performed by him. To feature thereto, we'd prefer to introduce a mechanism that albeit the person goes out of frame for a bit whereas and once more returns, he wouldn't be thought to be a replacement person and also the time length calculation for him would continue from once he left the field of read of the camera. This may be achieved by configuring the system to pause the timer for a selected person for a few time even once he has gone out of the frame. The timer resumes if the person returns among the field of read of the camera among the stipulated time length that the timer has been paused for a selected person. The length that the timer must be paused would be predetermined and pre-determined once observant the patterns of many police investigation videos. If the person is out of frame for a extended length than what was predetermined, the timer eventually resets for that person. This way, the field of read of the police investigation camera wouldn't be a bar in detection method to some extent.

The probability of chase errors will be reduced by employing a additional advanced motion model like constant acceleration, or victimization multiple Kalman filters for each object. In gift implementation, it's been assumed that each one individuals move during a line with constant speed. We'd prefer to build our trained classifier model additional sturdy by coaching it with additional forms of violent activities in order that in order that error is reduced.

Besides the prevailing suspicious and violent activity detection system, AN identity detection system (in the shape of face cataloguer) can even be developed by victimization the PTZ (pan-tilt-zoom) camera. It will be accustomed acquire high-resolution pictures of human faces concerned in violent or suspicious activity as detected by the system. There would be AN underlying face detection rule incorporated within the system. The instant any suspicious or violent activity is detected, the face recognition module are going to be triggered and also the PTZ camera would consequently be controlled to amass a zoomed-in image of the human faces concerned within the abnormal activity. This information therefore non inheritable will later be used for investigation functions. The practicality is like once the (frontal) face image is detected, the camera gets centred on the face and is zoomed-in. The pan and tilt of the camera gets adjusted supported the relative displacement of the centre of the face with relevance the centre of the image. With the assistance of intrinsic standardisation parameters of the camera and also the current zoom level, the relative image coordinate displacements square measure translated into desired or relative pan and tilt angles.

## V. Conclusion

The good police investigation systems will remodel the normal video police investigation systems from mere information acquisition tool to the data and intelligence acquisition systems. The situational info captured together with real time acquisition of the police investigation information will give the system a very important mechanism to keep up the situational awareness and conveyance higher security. In our technique, we've got used the HOG (histogram of familiarised gradients) options of the frames to coach the SVM classifier to classify the input check frames to either traditional frame or frame showing some act of human violence. If it's of the second sort, then associate degree alarm is raised by the system to alert the observer sitting within the room concerning the violent activity. Our system may also perform the motion detection of a person's victimisation the mixture of mathematician model, followed by the morphological operations to scale back the noise from the foreground detected binary image, so blob analysis to find the connected parts. The system keeps track of the time length that a specific person is found getting an area being monitored. The system alerts the observer within the room indicating that suspicious activity is also distributed by the person as he has been found loitering by the system for over a predetermined time length. We've got created it to figure with real time video acquisition and creating it a lot of sturdy by victimisation a lot of range of relevant

options excluding the HOG options to coach the SVM to coach to reduce the quantity of false alarms. We are going to attempt to upgrade the system to figure with inputs from the multiple cameras placed at unbroken angles of an equivalent region of interest to form the prediction a lot of correct. The feature vectors generated ought to be optimized to scale back their dimension so as to scale back dallying within the operation.

## References

- [1] Spasic, N.: Anomaly Detection and Prediction of Human Actions in a Video Surveillance Environment, Master Thesis, Computer Science Dept., Cape Town University, South Africa, December (2007)
- [2] Haque, M., Murshed, M.: Robust background subtraction based on perceptual mixture of gaussians with dynamic adaptation speed. In: IEEE International Conference on Multimedia and Expo Workshops (ICMEW 2012), Melbourne, Australia, pp. 396–401, July 2012
- [3] Regazzoni, C., Cavallaro, A., Wu, Y., Konrad, J., Hampapur, A.: Video analytics for surveillance: theory and practice. *IEEE Signal Process. Mag.* 27(5), 16–17 (2010)
- [4] Reddy, V., Sanderson, C., Sanin, A., Lovell, B.C.: MRF-based background initialisation for improved foreground detection in cluttered surveillance videos. In: 10th Asian Conference on Computer Vision (ACCV 2010), Queenstown, New Zealand, Vol. Part III, pp. 547–559, November 2010
- [5] Liu, L., Tao, W., Liu, J., Tian, J.: A variational model and graph cut optimization for interactive foreground extraction. *Signal Process.* 91(5) (2011)
- [6] Suau, X., Casas, J.R., Ruiz-Hidalgo, J.: Multi-resolution illumination compensation for foreground extraction. In: 16th IEEE International Conference on Image Processing (ICIP 2009), pp. 3189–3192, November 2009
- [7] Loy, C.C.: Activity understanding and unusual event detection in surveillance videos, Ph.D. dissertation, Queen Mary University of London (2010)
- [8] Cavallaro, A., Steiger, O., Ebrahimi, T.: Tracking video objects in cluttered background. *IEEE Trans. Circuits Syst. Video Technol.* 15(4), 575–584 (2005)
- [9] Karasulu, B.: Review and evaluation of wellknown methods for moving object detection and tracking in videos. *J. Aeronaut. Space Technol.* 4(4), 11–22 (2010)
- [10] Cilla, R., Patricio, M.A., Berlanga, A., Molina, J.M.: Human action recognition with sparse classification and multiple-view learning. *Expert Syst. J.* (2013). Wiley Publishing Ltd
- [11] Javed, O., Shah, M.: Tracking and object classification for automated surveillance. In: Proceedings of the 7th European Conference on Computer Vision, Part-IV, pp. 343–357 (2002)
- [12] Li, T., Chang, H., Wang, M., Ni, B., Hong, R., Yan, S.: Crowded scene analysis: a survey. *IEEE Trans. Circuits Syst. Video Technol.* 25(3), 367–386 (2015)
- [13] Patino, L., Ferryman, J., Beleznai, C.: Abnormal behaviour detection on queue analysis from stereo cameras. In: Advanced Video and Signal Based Surveillance (AVSS), 12th IEEE International Conference, pp. 1–6, August 2015
- [14] Chan, A.B., Morrow, M., Vasconcelos, N.: Analysis of crowded scenes using holistic properties. In: Performance Evaluation of Tracking and Surveillance workshop at CVPR (2009)
- [15] Lee, L., Romano, R., Stein, G.: Introduction to the special section on video surveillance. *IEEE Trans. Pattern Anal. Mach. Intell.* 8, 740–745 (2000)
- [16] Zhan, B., Monekosso, D.N., Remagnino, P., Velastin, S.A., Xu, L.Q.: Crowd analysis: a survey. *Mach. Vis. Appl.* 19(5–6), 345–357 (2008)



- [17] Baumann, A., Boltz, M., Ebling, J., Koenig, M., Loos, H., Merkel, M., Niem, W., Warzelhan, J., Yu, J.: A review and comparison of measures for automatic video surveillance systems. *EURASIP J. Image Video Process.*, 1–30 (2008)
- [18] Cai, Y., de Freitas, N., Little, J.J.: Robust visual tracking for multiple targets. In: *European Conference on Computer Vision, LNCS, Vol. 3954*, pp. 107–118 (2006)
- [19] Chang, T., Gong, S., Ong, E.: Tracking multiple people under occlusion using multiple cameras. In: *British Machine Vision Conference*, pp. 566–575 (2000)
- [20] Donalek, C.: *Supervised and Unsupervised learning* (2011)
- [21] Benezeth, Y., Jodoin, P.M., Saligrama, V.: Abnormality detection using low-level co-occurring events. *Pattern Recogn. Lett.*32(3), 423–431 (2011)
- [22] Isupova, O., Kuzin, D., Mihaylova, L.: Abnormal behaviour detection in video using topic modeling. In: *USES Conference Proceedings. The University of Sheffield, June 2015*
- [23] Cosar, S., Donatiello, G., Bogorny, V., Garate, C., Alvares, L.O., Bremond, F.: Towards abnormal trajectory and event detection in video surveillance
- [24] Morphology. <http://homepages.inf.ed.ac.uk/rbf/HIPR2/morops.htm>
- [25] How a Kalman filter works, in pictures. <http://www.bzarg.com/p/how-a-kalman-filter-works-in-pictures/>
- [26] Introduction to support vector machines. [http://docs.opencv.org/2.4/doc/tutorials/ml/introduction\\_to\\_svm/introduction\\_to\\_svm.html](http://docs.opencv.org/2.4/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html)
- [27] Background subtraction for detection of moving objects. <https://computation.llnl.gov/casc/sapphire/background/background.html>
- [28] Machine Learning, Part I: Supervised and Unsupervised Learning. [http://www.aihorizon.com/essays/generalai/supervised\\_unsupervised\\_machine\\_learning.htm](http://www.aihorizon.com/essays/generalai/supervised_unsupervised_machine_learning.htm)
- [29] BLOB Analysis (Introduction to Video and Image Processing) Part 1. <http://what-when-how.com/introduction-to-video-and-image-processing/blob-analysis-introduction-to-video-and-image-processing-part-1/>