ISSN 2395-1621

Suspicious Activity Detection using Haar Cascade Algorithm

[#1]Sagar Mujumale, [#2]Pallavi Ohol, [#3]Akshita Shewale, [#4]Rutuja Balkundre, [#5]Prof. Aparna Hambarde

[1]mujumalesagar@gmail.com
[2]pallaviohol694@gmail.com
[3]akshita.shewale26@gmail.com
[4]rutuja12balkundre@gmail.com

[#12345]Department of Computer Engineering,
KJCOEMR, Maharashtra, India.

ABSTRACT

Detection of Suspicious Movement or Activities had attracted an increasing level of attention. Automated offline video processing systems have been used for various purpose such as a proof of a crime or in a forensic labs. However, very little success has been achieved on this matter. In this paper we are introducing a system that works for surveillance systems installed in indoor environments like entrances/exits of buildings, corridors, High Sensitive area etc. We are proposing a framework that processes video data obtained from a CCTV camera fixed at a particular location. First we are obtaining the foreground objects and subtracting that foreground image from background subtraction. After subtracting this objects are then classified into people and inanimate objects Using Haar like features for face, head and people Detection in video sequences. Algorithm for detection of group of peoples suspicious movements and detection of Single persons suspicions actions and behavior. We track those objects by using Real-Time blob matching algorithm. We are also adding the fire detection facility into the proposed system. Hence we can know the cause of Fire.

Keywords: object detection, image processing, background subtraction, tracking, surveillance, Blob matching, Haar, fainting, fire detection.

ARTICLE INFO

Article History

Received: 27th May 2018

Received in revised form : 27th May 2018

Accepted: 31st May 2018

Published online : 1st June 2018

I. INTRODUCTION

Video Surveillance System is a collection of various types electronic and wireless components as well as video. It records videos of the various important locations for monitoring purpose. Due to the increasing rate of the crime around the world, many organizations are deploying Video Surveillance System at their locations for monitoring purpose. Increasingly, police and security staff relay on Video Surveillance System to facilitate their work. This practice is most evident in large public transportation areas such as metro stations, Airport as well as highly sensitive areas. However these systems remain highly insensible and the person monitoring the video displays find it extremely difficult to be attentive to randomly occurring incidence as the concentration of an operator will reduce drastically as time passes. Although automated Video Surveillance System do exist, they have been used mainly for offline videos analysis after an event has occurred, most notably in the case of riot investigations and forensics. At present, these surveillance systems are of magical help for real-time alerts. The functions of a automated surveillance system is to attract the attention of monitoring human Operator to the occurrence of a user defined suspicious behaviour when it happens. Mainly there are two challenges that stand in the face of developing fully automated behaviour recognition. First, objects of interest, such as people and luggage instable means of describing events must be found. The majority of researches till now day have involve machine learning to detect suspicious behaviour. To our knowledge, we purpose here a complete semantic based solution to the behaviour detection problem that addresses the whole Process from pixel to behaviour level. Here, we are using Haar like Feature algorithm for finding the suspicious activity in crowded area or group of people and to find the single persons suspicious activity and its behaviour. To find single person's suspicious activity we are using Haar like algorithm in our proposed system. It is mainly used for adjusting the image intensities to enhance the contrast. To find the suspicious activity in crowded area or a group of

© 2018, IERJ All Rights Reserved

peoples we are also use a Haar algorithm. In most of the previous group detection activity has used Hidden Markov Model which is used for the fixed number of peoples in a group for detecting the suspicious activity where the number of peoples in a group may vary (i.e. the number of people can leave or add in the group)which is our daily life scenario in a public transport area or in crowded areas , while grad algorithm allowed us to detect the suspicious activity for any number of peoples in a group even after the variation into the group of people. Also we have also add a feature known as Fire Detection which comes in handy to know the actual Cause of the fire in sensitive areas and could be avoided in the future.

II. EXISTING SYSTEM

Many researchers has been worked in this area. Most of the work is focused on machine learning for training the system to identify various anomalous scenario's. Recognition of human behaviour is a vast term which covers different categories and activities.[7] for example, crowd recognition, it should recognize the behaviour or movement of crowd or group of people rather than focusing on individual behaviour. On the other hand other human activities like[8]stretching, exercises and other gestures are[9] relatively simpler and periodic. These are different human nature and therefore require[5] different detection techniques with the help of body models and[6]space-time space. Some work also made use of the stereo cameras to model the objects that are present in the scene in 3-D. Elhamod and Levine [1] proposed a system for detecting suspicious activities in public transport Area. 3-D object -level information is obtained by detecting and tracking object/people using blob matching technique. Colour histograms are used for blob matching technique. This paper is focuses on automatically flagging suspicious behaviour in public areas or sensitive areas. For example[5] loitering [6] [10] abandoned objects and fighting. This behaviour may occur after a significant period of time. Sometimes they involve more than one object therefore such matters as finding identity of tracking persona and object classification must be addressed.[1] Elhmod and Levine introduced an approach that detect semantic behaviour based on object sand inter-objects motion features. kim at el deals with the detecting and tracking moving objects through single camera. The proposed system used RGB colour background modelling to extract Moving regions. Blob labelling is often used to group moving object. Some work [3] [4] also involves using a stereo camera. A stereo camera is a set of two camera mounted adjacent to each other, but separated by a small distance. These cameras face in same directions, but the images captured by them differs slightly due to the spatial difference between the lenses of both cameras. These images can be processed together to get 3-D locations of the objects. Such types of camera allowed for good occlusion detection and removal. But the Computation cost increases as we need to process two similar images for detecting suspicious activity more work has to be done in a machine learning area. Elhamod etal has used simple approach called as semantic-based approach. This allow us to define Suspicious activities based on human reasoning and logic. This eliminates the use of training the system, as it done in machine learning program, for identifying suspicious activities. This approach gives better result than machine learning as latter depends on a standard reliable dataset for training. Thus it makes Machine Learning approach less reliable. Therefore we are using a semantic-based approach.

III. PROPOSED SYSTEM

We are proposing a system which is able to process off-line as well online video surveillance for detecting suspicious activities based on a Human behaviour. We have also added a extra feature in our system known as fire detection. Here we are using semantic knowledge to detect the suspicious behaviour of single person as well as behaviour of number of peoples(i.e. group of peoples)like walking together, fighting in a group, terrorist group attack etc. We are using a monitor displays which are connected via internet to the cameras for monitoring the activities, Once it has captured a any suspicious behaviour of human or any lathering object it will take a snapshot of a that activity and send it to the user(i.e. monitor display) as well as it will generate the Alert at the same time for every suspicious activity. Same logic is used for the fire detection system. Whenever there is a smoke or fire it will take its snapshot and will generate the Alert for the person which is monitoring the display.

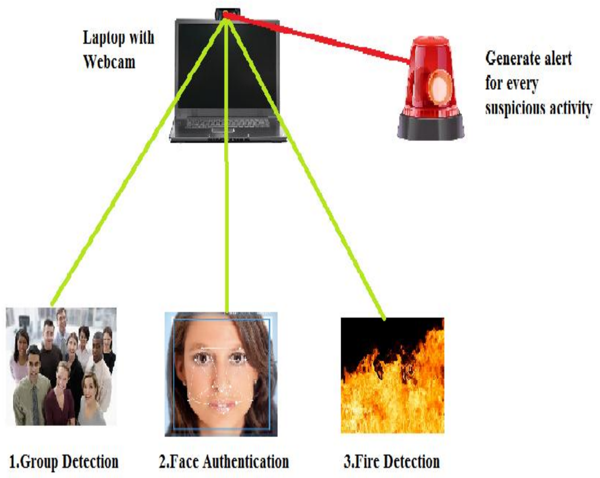

Fig. Proposed System

IV. ARCHITECTURE DIAGRAM

A. Find the Suspicious Activity

In this section we are finding the type of different suspicious activities like:

i. Lay down body

Which means if there is any human body in a 90 degree position on a ground (i.e. activity like fainting) it will go into the categories of lay down body.

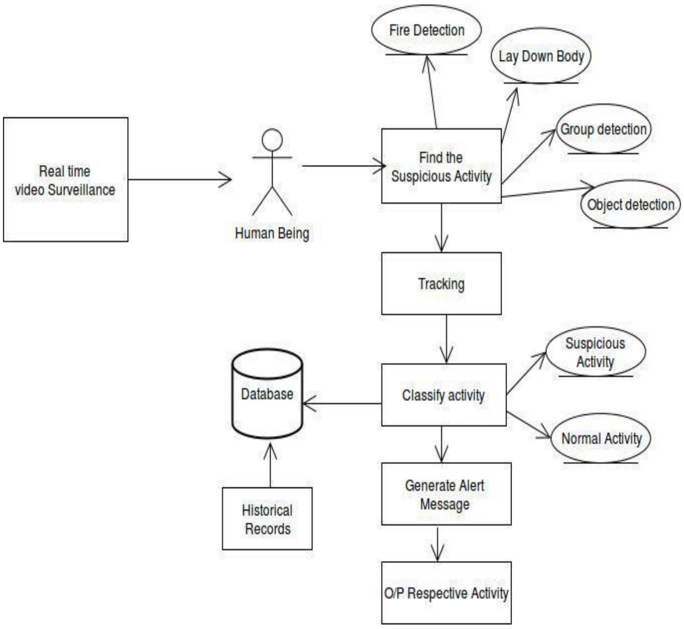

Fig. Architecture Diagram

ii. Group Detection:

If there is any suspicious activity in group like walking together, or fighting in a group or any terrorist activities perform in a group then it will go in categories of group detection and generate alert indicating suspicious activity related to Group Detection.

iii. Object detection:

If there is any abandoned object(i.e. luggage ,purse, sack bags etc) found on a same place even after threshold value has pass then system will generate a Alert message under the categories of Object detection. Same logic goes for loitering objects too.

iv. Fire Detection

if there is a smoke or fire(i.e. suddenly rise in a temperature) then system will take its snapshot and will send it to the human operator which is monitoring that system and will generate the Alert under the categories of Fire detection, which is also useful in knowing the cause of fire in the area.

B. Classify Activity

After capturing the image it processes the image then decide whether it is a suspicious activity or a Normal Activity based on the Semantic knowledge. it processed the image from the database. It also processed the images from the Historical Records.

C. Generate Alert message

After classifying stage if the Activity is classify as Suspicious then it will Generate a alert message related to the categories of suspicious activity and will send to human operator monitoring the display monitors.

D. Output Respective Activity

After generating the alert message it will send the message to the human operator that is monitoring the display which are connected via internet to the display monitors it will also specify the categories of activity like whether it is lay down body or it is group detection or it is Object Detection or fire Detection.

I. Haar like Algorithm

It is analyzing the video content at real-time frame rate and with a false alarms detection rate as small as possible. This algorithm can be dedicated and specifically parameterized in certain application and restrained Environment. They also face on many challenging issues in order to correctly analyze human activities. Main attention is focused in detecting face, head and people body in videos. Faces can be seen as small objects when people are distant from the camera.

1) Face detection using couple cell response features -

A couple cell feature consists of two adjacent cells c = {1, 2} forming a ells. The couple can be oriented horizontally, vertically or in diagonal. Pixel distributions v1 and v2 of the two cells c1 and c2 respectively are compared to give the feature response.

Acell pixel distribution is represented by an average E (the expectation term) ofthe pixel grey levels within this cell. Dimension X = {x, y,w, h}the magnitude and sign of the difference Δ between these two mean values:

$\Delta = {}_c(I(x))$ ---------------------------------(1)

$\Delta = v2 - v1$ ----------------------------------(2)

$= b \ \mathrm{sign}(\Delta)$, $2b < \Delta < 2b+1$ -----------------(3)

The CRR feature is of dimension 17 and its features represent varying ranges of pixel distribution differences:

(b; Δ) = (−8; [−255 :−128[), (−7; [−127 : −64[)...(0; [−1 : 1[), (4)

(1; [1 : 2[), (2; [2 : 4[)...(8; [128 : 255]) ----------(5)

© 2018, IERJ All Rights Reserved

The couple cell feature is calculated across an image at various scales byvarying the cell width w and height h as follows:

(w, h) = {(1, 1)(1, 2)..(1, 8), (2, 1)..(8, 8)} (6)

Adaboost training:

feature position = x

dimension width = w , height= h

That we encapsulate into X= {x,w, h}, a weight probability distribution is calculated in equation from all samples indexed by i.

The positive (object) samples defined by = 0 and labelled(I) = 0

The negative (non-object) samples defined by ¸ = 1 and labelled(I) = 1.

$$G^{(t)}_{\chi,\lambda}(b) = \sum_i w(i)\ \delta_\chi(i,b)\ (1 - |c(i) - \lambda|)$$

$$\delta_\chi(i,b) = \begin{cases} 1 & \text{if} \quad f_\chi(i) = b \\ 0 & \text{else} \end{cases}$$

It classifies samples according to their feature response and their weight:

Feature described above.

The initial sample weights w(i) are normalized over the distribution of samples ∑ ( ) = 1 A classifier *h* (i) at iteration t is associated with each feature position X in the samples template area. It classifies samples according to their feature response and their weight

$$h^{(t)}_\chi(i) = \begin{cases} 1 & \text{if} \quad G^{(t)}_{\chi,0}(f_\chi(i)) > G^{(t)}_{\chi,1}(f_\chi(i)) \\ 0 & \text{else} \end{cases}$$

The Adaboost weak classifier h(i) at iteration t is chosen at the feature position where the classifier is the most discriminative among all other classifiers

$$\chi_m = \mathrm{argmin}_\chi \{e_\chi\}$$

$$e_\chi = \sum_i w(i)\ |h^{(t)}_\chi(i) - c(i)|$$

In this equation the error term represents the sum of sample weights of the badly classified samples at feature location X. Sample weights are then updated in the following equation where Z represents a normalizing factor:

$$\alpha^{(t)} = \frac{1}{2} ln \frac{1 - e_{\chi_m}}{e_{\chi_m}}$$

$$w(i) \longleftarrow \frac{w(i)}{Z} \begin{cases} e^{-\alpha^{(t)}} & \text{if } h^{(t)}_{\chi_m}(i) = c(i) \\ e^{+\alpha^{(t)}} & \text{else} \end{cases}$$

Objects are detected as positive candidate using the trained data

$$H(i) = \sum_t \alpha^{(t)} h^{(t)}_{\chi_m}(i) > 0$$

V. RESULTS AND DISSCUSION

The evaluation of activity recognition has been challenged by number of different difficulties. Firstly measures of activities of the interest are of high complexity, which has become a issue in the presence of clutter in the test scenario. Second issue is the delinquency of professional and challenging high quality data sets currently available for testing. The definition of suspicious activity is dependent on its Scenario. For example person waiting for a bus for 15 minute is normal; whereas for ATM, a person waiting for such a long time is highly suspicious. These challenges lead to the inconsistencies among the experimental results in different literature papers. We have used carefully selected standard public data sets to test the proposed framework. These data sets are CAVIAR (PETS 2004) and PETS 2006. The major limiting factor in the field's Development is a trend toward focusing on non-professional data sets.

A. By using proposed algorithm we can track multiple peoples simultaneously. In the System two outgoing people are tracked. The bounding box is drawn around the each person. Every person is given a unique ID. In all the subsequent frames , the detected objects are compared with objects in previous frames. if an object in a current frame is found match to previous frame, it is assigned the same ID.

VI. DETECTION OF SUSPICIOUS ACTIVITY VARIOUS CASES

Case 1: Loitering at an ATM

Identifying loitering at an ATM is crucial from a security point of view. Today though the ATM provides lot of different services to the users; though they are still predominately used for their basic function of withdrawing money. The activity of withdrawing money is usually takes around 50-60 seconds under normal circumstances.

Case 2: Loitering at Bus Station

As the bus station is a public transport area, identification of the loitering become complicated. There is always lot of different activities going around the bus station. A person can be wait for bus for any length of time. Hence it is nearly impossible to identify a genuine loitering case from a false one without human innervations. Thus here we are considering all loitering cases are genuine, and the system mark the person as a loitering when that person is in the frame more than define time threshold. Then it is up to a human operator to decide whether to take further action or not. Identification of the abandoned luggage is crucial from security point of view as the abandoned luggage item may contain explosive. The luggage item is marked as abandoned if its owner goes away leaving the luggage Item behind.

© 2018, IERJ All Rights Reserved

Case 3 : Fire Detection

In public or private places there are some cases whereas suddenly that place is sets on fire

and we don't know the cause of it. For example in a bus station or in shopping malls there are cases of suddenly in particular areas of shopping malls are on fire and the staff and mall manager does not know the cause of it.

VII. FUTURE SCOPE

The system works mainly in indoor environment like entrance or exits of the building, corridors etc. The orientation and location is a important feature of a camera. The proper position to capture perfect human body shape would be to directly place the camera in front. But it is not possible in real-life scenario's. Hence the camera is placed overhead in a downward angle. The exact degree of inclination is not important. Not as an independent document. Please do not revise any of the current designations.

VIII. CONCLUSON

In this research work, a complete semantic-based simple approach that depends on object tracking has been introduced. Our approach begins with translating the object obtained by background segmentation into semantic entities in the scene. These objects are tracked in 2-D and classified either being animated or inanimate. Then, their motion features are calculated and recorded in the form of historical records. Finally, behaviors are semantically define and detected by continuously checking those records against predefined rules and conditions. This approach ensures Real-Time performance, adaptability, robustness against clutter and camera non-linearity, ease of interfacing with human operators, and elimination of the training required by machine-learning-based methods.

REFERENCES

[1] N. Amani, A. Shahbahrami and M. Nahvi, “A New Approach for Face Image Enhancement and Recognition”, International Journal of Advanced Science and Technology, vol. 52, (2013).

[2] M. Turk and A. Pentland, “Eigenfaces for recognition. Journal of Cognitive Neuroscience”, vol. 3, no. 1, (1991).

[3] A. J. Colmenarez and T. S. Huang, “Face Detection and Recognition. Berlin: Springer-Verlag”, Series F: Computer and Systems Sciences, vol. 163, (1998), pp. 174-185.

[4] G. Wang and Z. Ou, “Face Recognition Based on Image Enhancement and Gabor Features”, Proceedings 6th World Congress on Intelligent Control and Automation, Dalian, China, (2006) June, pp. 9761-9764.

[5] K. Kim, Editor, “Face recognition using principle component analysis”, International Conference on Computer Vision and Pattern Recognition, USA, (1996).

[6] H. Yu and J. Yang, “A direct LDA algorithm for high dimensional data-with application to face recognition”, The Journal of Pattern Recognition, vol. 34, (2001), pp. 2067-2070.

[7] P. Penev and J. Atick, “Local Feature Analysis: A General Statistical Theory for Object Representation”, Network Computation in Neural Systems, vol. 7, (1996).

[8] X. He, S. Yan, Y. Hu, P. Niyogi and H. -J. Zhang, “Face recognition using Laplacianfaces. IEEE Trans. on PAMI, vol. 27, no. 3, (2005).

[9] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman, “Eigenfaces versus fisherfaces: Recognition using class specific linear projection”, IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 7, (1997).

© 2018, IERJ All Rights Reserved