Real Time Motion Detection Video Surveillance For Categorizing Between Human and Non-Human

Meghna Jain, Kriti Agarwal, Vrajesh Sheth,Vaishali Gaikwad(Mohite)
Dept. of Information Technology,
K.J.Somaiya Institute of Engineering & Information Technology,
Mumbai (India)
meghana.jain@somaiya.edu,kriti.a@somaiya.edu,vrajesh.s@somaiya.edu,mohite@somaiya.edu

Abstract— An Automated Video Surveillance and monitoring system has a rich history. Traditional video surveillance systems are large in number but still retrenched in an extensive manner. The system targets at tracking an object and segregates it as Human or Non-Human entities, wherein the non-human entities would be further analysed into its respective categories which would help in subsequent analytics.The system when recognizes suspicious activity, it is captured instantly and the alarm is triggered for the security purpose. It is necessary to introduce an application that automatically formulates the images captured in order to detect precarious situations or undesirable encroached objects. Object detection is a mandatory step in automated video surveillance. Foreground extraction in harmony with background subtraction is further collaborated with the thresholded image for revelation of entities. The system engages a contemporary combination of Background Modelling, Support vector machine (SVM) and a Human Detection for Surveillance (HDS) System. The HDS system assimilates a Histogram of Oriented Gradients based on a human detector which is in limelight for its performance in detecting humanistic appearances. Detailed analysis is carried out on the performance of the system on various test videos.

IndexTerms—Video surveillance system, Tracking, Background subtraction, Producing alarm, Human Detection for Surveillance system (HDS), Support Vector System(SVM).

I.INTRODUCTION

Objects Detection and Image Compression are the most important and challenging fundamental tasks of computer vision applications, like surveillance, monitoring, robot technology, object recognition, authorization etc. Object recognition should be invariant to view point changes and object transformations and also robust to noise and occlusions. It is still an open problem due to the variety and complexity of object classes and backgrounds. Numerous approaches have been put forward for object detection from videos, mainly dedicated to human monitoring and visual surveillance. Although the specific requirements vary between surveillance systems, all are facing numerous problems. Usually, an operant concerns him only with the detection of objects whereby being ignorant of the other factors. Moving objects and still images are supposed to go hand in hand, but some systems consider only one of the two while keeping the other one untouched. There are a number of approaches available out of which background subtraction seems to be the most promising prospect. As soon as the image is captured it will be fetched to a standard format which aims to describe the procedure of standardizing it. Human Detection for Surveillance (HDS) is used for detecting humans and non-human entities and Support vector machine helps you identify and classify the non-human entities into its further categories.

This paperis organized as follows. Section II describes the related methods available. Section III briefly describes the proposed methodology. Section IV deals with the comparison between the existing system and the proposed system. Section V includes the conclusion and future enhancement.

II.RELEVANT WORK

An outstretched research was done over the techniques, methodologies and algorithms accordant to motion detection and categorizing between human and non-human. After doing a thorough survey over techniques, methods and algorithms which are related to motion detection, object recognition, more importantly related to detecting of moving objects during a video/frames. If suitable methods are used it would rule out the practical problems related to shadow and illuminous changes.
Many researchers have put forward their understandings and research's regarding motion detection under indoor as well as outdoor scenes helping with solutions to above problems.

In [1] proposed that camera consumes large storage space if it records every bit and results in limitation of storage space and duration of video. It is practically impossible for human to review each and every thing that gets recorded in a video and therefore these disadvantages are solved by capturing only images using motion detection algorithm i.e capture those images which contain motion or any changes

. In [2],The system uses Adaptive Background Modeling Algorithm which is based on the Gaussian Mixture Model. The system incorporates a human detector which known for its performance in detecting humans in still images.
[3] proposed a temporal differencing to detect the moving object and trigger alarms when necessary giving high accuracy. It helps in high performance and better speed.
[5] Robustness against factors like deformed and changeable shape was handled by the technique of motion detection and object tracking which is shale and rotationally invariant also providing faster processing time.
[6]Modelling activities provides statistical activity recognition using temporal differencing.
[7] Reduction of unwanted recording of videos was done using motion detection approach. It consumes low power. Background subtraction technique is mainly used to detect moving entities and also helps in reduction in illuminous changes.
[8] Noise was maintained by using motion detection method which scans from top to bottom for detecting presence of entities around.
Background modelling is the best method which is able to detect foreground objects against new backgrounds.

III.OVERVIEW OF THE PROPOSED WORK

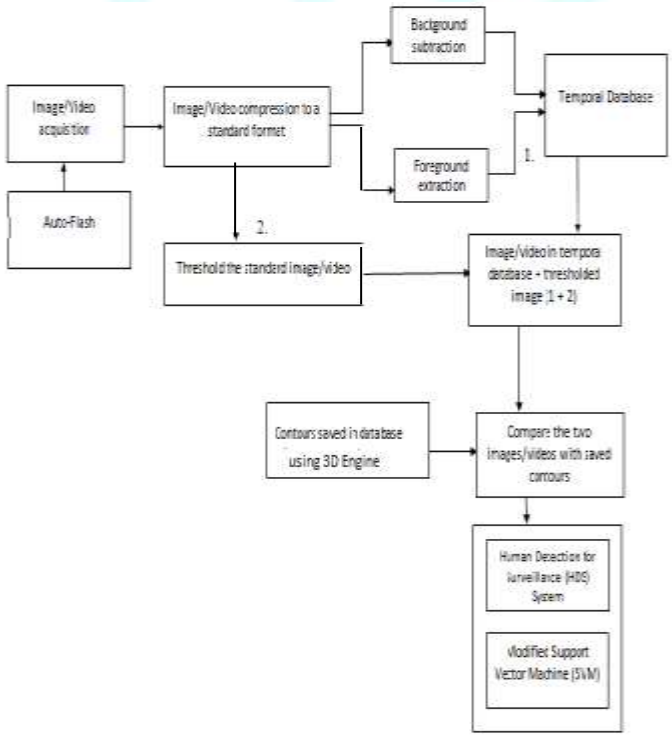

Fig .01 – Block Diagram of proposed system

The proposed model is depicted in fig 01. The video processing techniques that are used in the proposed system has been explained in the following sections.
The overview of the proposed model is as follows:

1. Acquiring the live video feed into the webcam is the fundamental step in video surveillance. Video processing is done contiguously. So a video sequence is composed of series of frames.

2. While the video is being acquired, due to ill light conditions auto flash would help enhance the illumination to get a better and clear image.

3. In the process of capturing the live video, the acquired images will be first brought to a standard format.

4.Background subtraction is bestowed to the standard frame leading to foreground extraction.

5. This extracted frame will be saved in the temporal database.

6. The standard frame will be taken again and thresholded for extracting the background entities as well.

7.The standard frame and the thresholded frame will then be coalesced for further analysed, we can compare these frames captured with saved contours of objects also these frames will be compared to the previous frame to detect the motion.

8.Human Detection for Surveillance (HDS) System will categorise the entities into human and non-human respectively.

Support vector machine (SVM) further identifies and classifies the non-human entities into its respective categories. These two modules will be combined together as single system for faster processing speed.

The entire system will be developed using OpenCv which is an open source library for real time computer vision.

Activity Behaviour of the human is scrutinised in the background modelling module. If the abnormal behaviour is tracked in the scene, the system automatically takes the snapshot of the detected entity and executes the alarm according to the user settings.

The proposed model is depicted in fig 01. The video processing techniques that are used in the proposed system has been explained in the following sections.

A. *A. Video Pre-Processing:*

The Videos captured by the Surveillance System are divided into frames and further pre processing is carried out to on the frames to enhance the contrast of the frames or removal of noise, and colour conversion of the image frames. This process improves the video/frames quality by suppressing unwanted distortion. Also to avoid unwanted result due to illumination changes auto flash system is provided.

B. Image Formatting

The frames evolved from the captured video can be of any format supporting it resulting in different quality and more space for storage. To avoid this difference we get the images in frames into a standard format without compromising with its output quality and saving space irrespective of input video quality. GIF, PNG, JPEG, and TIFF files are designed for different graphic needs. There are chances that we need to use the combination of all the provided formats depending the task on hand and the current image available in frame.

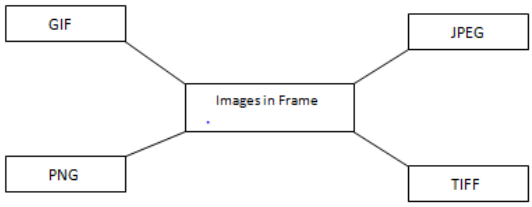

Fig02.Image Formatting Details

C. Background Subtraction Modeling

Identifying objects from a video sequence is a fundamental and critical task. Background Subtractor MOG and MOG2 will be used for detecting moving objects in videos with a static camera. This is used for generating Foreground Mask of the moving object in the current frame. Absolute Difference will be calculated by comparison of the current frame and the background frame which can be in general background given as the characteristics of the frame. This Foreground mask image will be stored in temporal database for further processing stage.

MOG and MOG2 are Gaussian Mixture Based adaptive Background Subtraction and Foreground Segmentation Algorithm. The MOG model is a method where each background pixel is model in a mixture of K Gaussian Distribution {k=3 to 5}. The MOG2 on the other hand selects an appropriate number of Gaussian distribution of each pixel of the background image. The MOG 2 provides better adaptability to the scenes changing due to illumination effects or due to any distortion in the image frame. In general illumination changes can be due to sudden change in the weather or bad day light or can be due to lighting the room. During such cases the objects from the scene can be removed out or moved from a place. In order to adapt from such changes, we choose a reasonable time period T and at time t we have XT = {x (t) , ..., x (t−T)}. For every new sample we update the training data set XT and re estimate pˆ(~x|XT , BG). However there are chances that the samples from recent history of frames may have values that belong to foreground objects. For such cases we should denote this estimate as p(~x (t) |XT , BG+FG). We use GMM with M components:

pˆ(~x|XT , BG+FG) = X M m=1 πˆmN (~x; bμ~m, σb 2 mI) …(1)

where bμ~ 1 , ..., bμ~M are the estimates of the means and σb1, ..., σbM are the estimates of the variances that describe the Gaussian components.

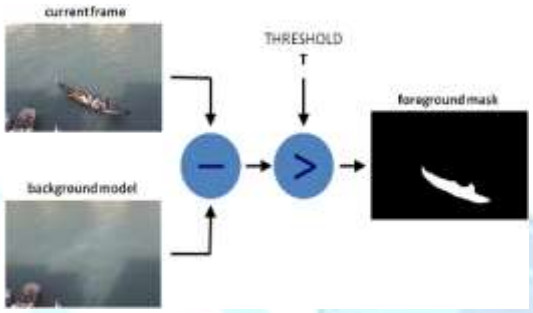

Fig3.Background Subtraction Modeling output

D. Thresholding

The improvised standard format image is again consider for detection or capturing of other background objects which were not considered in the previous stage of Backgroud Modeling using MOG and MOG2.This frame is saved in the temporal database along with the foreground mask image and are compared with the contours.

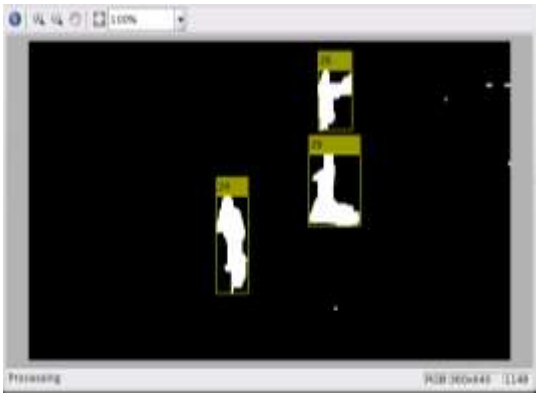

Fig.4.Threshoded Image capturing Background Motion

E. Contours and 3D Engine:

Since it's impossible to capture and detect all the objects in the frame, the proposed system will have all the predefined contours saved in the database. All these contours will be integrated with 3D Engine which will give an 360 degree rotated image of each and every object making its detection easier. This will also save the space in the Database.

Appearances of an object have a large number of variations due to photometric effects, scene clutter, and changes in shape, viewpoint changes. Different views of same object can give rise to widely different images. Different views of same object can give rise to widely different images. Hence 3D engine is required to understand the relationship between geometric shapes and their projections in any 3d entity. We must match an image to one of a huge number of possible objects, in any of the infinite number of possible positions.

.

F. HDS & Modified Support Vector Machine (SVM)

The saved objects from the frames are then processed further for Human Detection Surveillance System(HDS) that will help us to identify the entity if being human. Modified Support Vector Machine(SVM) being an binary classifier will identify the Non- human object and classify it further into category of Non Human entity. This will reduce the processing time of the captured image ,differentiating the objects in the single output frame.

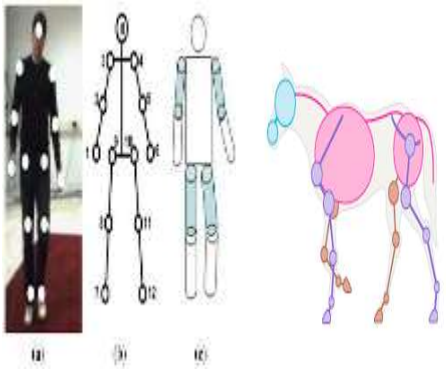

Fig.5:Humanand Non-Human Differentiating

G. Alarm Trigger

Since the System is Surveillance System the Alarm Triggering becomes the core module of this system. Whenever any unwanted or restricted entity or object is spotted it will be identified and the alarm will trigger. The Proper outcome of this step depends upon the outcome of the previous step i.e. HDS and Modified SVM.

IV.EXISTING SYSTEM V/S PROSED SYSTEM

In the existing system, the background image is subtracted from the current frame image and thresholded whereas in the proposed system the background is subtracted to extract foreground images and then thresholding is done further on the standard image so as to consider each and every suspicious object present in the frame. Since the existing system failed to recognize the objects in the background hence this laid a major loop hole in any real time detection system. Proposed system will be detecting foreground as well as background entities and classifying them in their respective categories. The static background subtraction is not resilient to illumination changes or long lasting changes in the scene therefore an auto flash is provided in the proposed system to get over the ill lit effects and enhance the illumination effects. Unlike other existing systems wherein two SVMs are used simultaneously we will combine SVM's as one SVM for human and non-human detection which will also classify its further categories to speed up the processing and detection time. In existing system, if robot is entering the video frame it is not possible to detect with the help of contours as it will be similar to the human contours hence as a future scope of the system infrared light can be inculcated which will detect the body temperature and will further help in categorizing it as human or robot. Open cv is used for our system which is an open source software hence cost expenditure is saved. Vision Library having a strong support and interface integrated with various platforms is used to implement Real time automated video surveillance.

V.CONCLUSION

The system targets at tracking an object and segregates it as Human or Non-Human entities, wherein the non-human entities would be further analysed into its respective categories which would help in subsequent analytics. The system engages a contemporary combination of Background Modelling, Support vector machine (SVM) and a Human Detection for Surveillance (HDS) System. The HDS system assimilates a Histogram of Oriented Gradients based on a human detector which is in limelight for its performance in detecting humanistic appearances. Appearances of an object have a large number of variations hence a 3d engine is used.. Detailed analysis is carried out on the performance of the system on various test videos.Furthermore, an option to take snaps and saving a predefine object contours in the system database for object detection and classification increasing the processing speed and reducing processing time with its implementation in OpenCv.

REFERENCES

[1] T.Deepika, Dr.P.Srinivasa Babu "Motion Detection in Real-Time Video Surveillance with Movement Frame Capture and Auto Record" in International Conference on Engineering Technology and Science-(ICETS'14) , Volume 3, Special Issue 1, January 2014.

[2] Sreedevi M, Yashwant Kumar Avuapati, Anjan Babu G, Sendhi Kumar R "Real Time Movement Detection for Human Recognition" in Proceedings of the World Congress on Engineering and Computer Science, 2012 Vol 1,WCECS 2012, October 24-26,2012.

[3] Prithviraj Banerjee and Somnath Sengupta " Human Motion Detection and Tracking for Video Surveillance",

[4] R.Shaalini & C.Shanmugam**,** R.Niranjan Kumar, G.Myilsamy, N.Manikandaprabu "Human Motion Detection and Tracking for Real-Time Security System" in International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 12, December 2013.

[5] Anand A.Dawood and Balasim A.Hussean "Motion Detection in Real-TimeVideo Streams using Moved Frame Background" in Journal of Babylon University in Pure and Applied Sciences,Volume 19,Issue 2011.

[6] Shih-Chia Huang "An Advanced Motion Detection Alogrithm with video Quality Analysis For Video Surveillance System" in the IEEE Transaction on Circuit and Systems for Video Technology, Vol 21, No1, Janaury 2011.

[7] . Asif Ansari, T.C.Manjunath., C.Ardil "Implementation of a Motion Detection System" in the International Journal of Electrical and Computer Engineering, 2008

[8] Robert T. Collins, Alan J. Lipton, and Takeo Kanade, "Introduction to the Special Edition on Video Surveillance",Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 22, No. 8, pp 745-757, August 2000.

[9] Ismail Haritaoglu, David Harwood and Larry S. Davis, "W4: Real-Time Surveillance of people and their Activities", Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 22, No. 8, pp 809-830, August 2000.

[10] N. Dalal and B. Triggs, "Histograms of oriented gradients for human.Detection", Proceedings of the Conference on Computer Vision and Pattern Recognition, San Diego, California, USA, pp. 886–893, 2005.

[11] Q. Zhu, S. Avidan, M-C Yeh, , K-W Cheng, "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients", Proceedings of the IEEE Computer Society Conference on Computer vision and Pattern Recognition, ISSN: 1063-6919, Volume 2, pp. 1491-1498, June 2006.

[12] M. Harville, G. Gordon, and J. Woodfill, "Foreground Segmentation using adaptive mixture models in color and depth", In Proceedings of the IEEE workshop on Detection and Recognition of Events in Video, 2001.