Abandoned Object Detection

Sagar
201601075
Indian Institute of Information Technology,Sri City
sagar.r16@iiits.in

***Abstract*— Due to the increase in terrorist activity in public places, security is now a major issue in public places. In this paper we have proposed an Abandoned Object Detection system which is one of the most hot and practically useful areas of computer vision. It has applications in automated video surveillance system for detecting suspicious objects. This can be used in crowded places like airport, railway station, etc. .An abandoned object is defined as one that has been lying stationary at a certain place with no apparent human attendance for an extended period of time. The first frame is taken as the background and all the corresponding frames are then subtracted from that frame. If a new object is found and that remains static at the same place for more than a certain number of frames the it is declared as abandoned object and an alarm is raised. The system allows the user to change the time after which the alarm is raised and the no. of times it can ignore the occlusion of object.**

Keywords*— *Abandoned Object,Background Subtraction,Blob Analysis

I. Introduction

Since last decade there had been rise in terrorist attacks on airports, train station, shopping malls, etc.. Despite deploying video surveillance systems, the surveillance prove to be inefficient since they require human manpower to monitor them. Hence there came out a need of automatic surveillance system which can identify abandoned object in the crowded places. These system can able to detect such object. Due to development of high processing cameras and need of security, abandoned object detection becomes the field of research in recent years. Problem of occlusion(the state when a person obstructs the view of the abandoned object) due to moving people make it complex, Other difficulties like high crowd, change in lighting condition and shadows makes this subject more complex.

In abandoned object detection we have to find out the objects that are being left alone by some person for sake of giving harm to society. In these paper we present an abandoned object detection algorithm which works efficiently on complex environment like crowd , shadow, Occlusion problem of people.

In this paper we have used background subtraction to detect any changes and then blob analysis to get the details for that change.

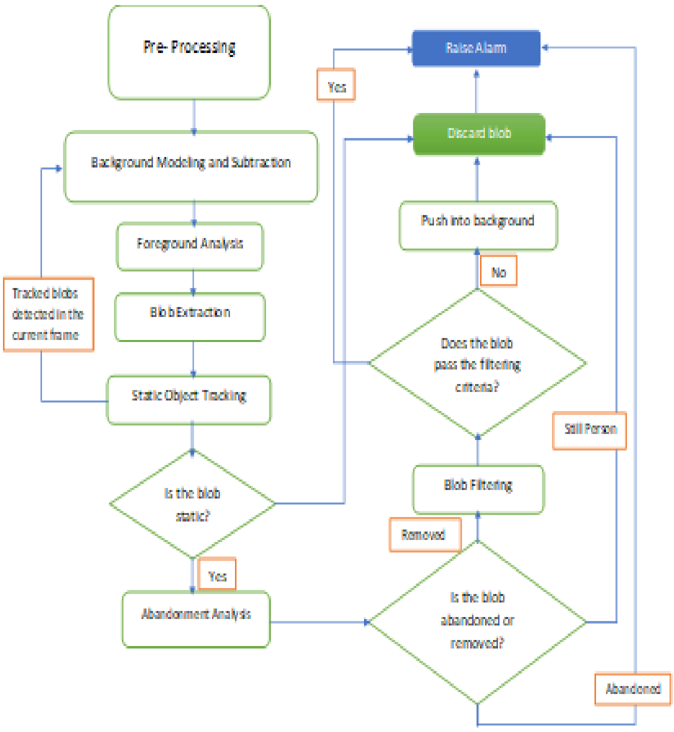
II. Proposed Methodology

Our aim is to differentiate between object belongs to foreground and background in video. The basic idea of implementation is process foreground and to find out static object which are not a part of background . After object detection raise the alarm if it is remain present for certain period of time. Flow chart in fig.1 shows the proposed algorithm.

Initially the region of interest is selected from the video. The complete surveillance will be done only in the region of interest specified. Then max number of abandoned objects to be detected and their minimum and maximum sizes are specified.

Now the frames of the video are traversed and are converted to YCbCr form to enhance the contrast. Now each frame is subtracted from the background frame. Then the subtracted image is converted to binary image. After that noise is removed from the binary image by joining the small balls. Now blob analysis is done to get the complete information about the blobs. Blob details like the area ,Centroid and the rectangle covering the blobs is received using blob analysis. Now in order to check if an object is abandoned or not analysis is done. The list of all the blob's centroid is traversed and the centroids are rounded off. After that it is added to a list of all the known centroids. If the centroid already exists in the list its hit count is incremented. Now after getting the complete data, if the hit count increases then the an alarm is raised. Or else due to occlusion if there is a miss then the miss count is incremented. If object appears within 7 frames then the miss count is brought back to zero.

III. Flow Chart



IV. Analysis and Development

- Background modeling and subtraction (BGS)
- Pre-processing
- Foreground analysis
- Blob extraction
- Blob tracking
- Abandonment analysis

A. *Background modeling and subtraction (BGS)*

- The first frame of the video is taken as the background.
- The background and the current frames are converted to YCbCr form.

BGS can be done by various algorithms, but here GMM is followed.

Gaussian mixture model: It assumes that the overall intensity at any pixel at each instant is produced by a combination of background and foreground processes, and each such process can be modeled by a single Gaussian probability distribution function.For each pixel in the current frame, the probability of observing the current intensity is given by:

$$P(X_t) = \sum_{i=1}^{K} \omega_{i,t} * \eta\left(X_t, \mu_{i,t}, \Sigma_{i,t}\right)$$

Here, K is the no. of distributions (K=3 here); $\omega_{i,t}$ is the weight associated with the i^{th} distribution at time t while $\mu_{i,t}$ is the mean and $\Sigma_{i,t}$ is the co-variance matrix of this distribution, η is the exponential Gaussian probability density function given by:

$$\eta(X_t, \mu_t, \Sigma_t) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma_t|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu_t)^T \Sigma_t^{-1}(X_t - \mu_t)}$$

Here, n is the dimensionality of each pixel's intensity value (e.g. n=1 for grayscale image and n=3 for RGB image). In order to avoid a costly matrix inversion and decrease computation cost, it is also assumed that the red, green and blue channels in the input images are not only independent but also have the same variance $\varsigma^2{}_{k,t}$, so the variance matrix becomes:

$$\Sigma_{k,t} = \sigma_{k,t}^2 \mathbf{I}$$

The K Gaussian distributions are always ordered in the decreasing order of their contribution to the current background model. This contribution is measured by the ratio ω/ς under the assumption that higher is the weight and lower is the variance of a distribution, more is the likelihood that it represents the background process.
For each pixel in an incoming frame, its intensity value is compared with the means of the existing distributions starting from the first one and a match is said to be obtained if its Euclidean distance from the mean is less than m standard deviations (m=3 is used here), i.e. it satisfies the following condition:

$$\left|I_t - \mu_{k,t}\right| < m * \sigma_{k,t}$$

Here, It is the pixel intensity while $\mu_{k,t}$ and $\varsigma_{k,t}$ are the mean and standard deviation of the k_{th} distribution at time t.
Since the background model is dynamic, it needs to be updated with each frame. While the weights are updated for all distributions, the mean and variance are updated only for the matched distributions. Following are the standard update equations used for this purpose:

$$\mu_t = (1-\rho)\mu_{t-1} + \rho X_t$$

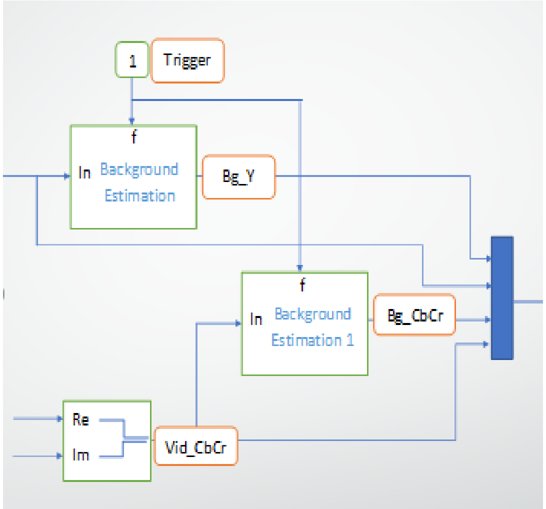
$$\sigma_t^2 = (1-\rho)\sigma_{t-1}^2 + \rho(X_t - \mu_t)^T(X_t - \mu_t)$$

$$\omega_{k,t} = (1-\alpha)\omega_{k,t-1} + \alpha\left(M_{k,t}\right)$$

Here X_t is the pixel intensity while $M_{k,t}$=1 for matched distributions and 0 for unmatched ones while ρ and α are learning rates. In the current work ρ and α are related as:

$$\rho_{k,t} = \frac{\alpha}{\omega_{k,t}}$$

Thus, while α is fixed for all distributions (α = 0.001 used here), ρ is smaller for higher weighted distributions. If none of the existing distributions match the current intensity, the least probable distribution (i.e. with the smallest value of ω/ς) is replaced by a new distribution with a high initial variance, low prior weight and the new intensity value as its mean.
This is the simplest way for background subtraction.



B. Pre-Processing

Pre-Processing involves 2 functions:

- Contrast Enhancement
- Noise Reduction

Contrast Enhancement: It improves the quality of low light video by normalizing the difference between the maximum & minimum intensities, using **YCbCr approach**. Convert RGB to YCbCr using "vision.ColorSpaceConverter" function. This let us convey the color information in a smaller amount of data than in the case of RGB. We perform YCbCr on the background and current frame.

Noise Reduction: This step reduces the white noise present in an input frame by smoothing the frame. It is particularly useful for low quality and is also needed to control the amount of noise that becomes visible in low light videos after applying contrast enhancement . This is accomplished by subjecting the image one of several smoothing filters including linear, Gaussian and median filters.

C. Foreground Analysis

Most existing BGS methods are far from perfect and tend to produce noisy outputs that need to be processed further before they can be used to extract useful objects. The noisy portions of the BGS output may contain false foregrounds produced for example during sudden lighting changes (a light is switched on or off) as well as actual foregrounds that are of no interest to further processing but can complicate it significantly. Foreground regions detected due to moving shadows are important examples of this latter category. Thus, this system has a separate foreground analysis stage that takes as its input the noisy foreground mask produced by the BGS module and removes all the false and uninteresting foreground pixels from it.

D. Blob Extraction

The refined foreground mask produced by the last step is subjected to a connected component detection algorithm to extract meaningful foreground objects while discarding any blobs that are smaller than a specified threshold. A simple, efficient but fairly accurate algorithm described in is used for this purpose. This algorithm uses a single pass over the image to detect external as well as internal contours and also labels the internal points of each connected component.

- This function **[Area, Centroid, BBox] = step(hBlob, Segmented)** gives area , centroid and Boundary box for a particular blob, discarding blobs that are smaller than a specified threshold.

E. Blob Tracking

In order to keep a track of all the blobs the complete data is appended in a matrix .The centroid is rounded off to the nearest multiple of 5.Then each time same centroid is found count is incremented.

If the centroid position and size(criteria) of the blobs are at that position already, then the Hit Rate of that blob is increased by 1.

F. Abandoned Object Analysis

- A blob with Hit Rate above certain threshold is discarded
- If the hit count increases a particular level, alarm is raised
- If any object is occluded, we will increase its Miss Rate by 1.
- If the miss count increases a particular level then the blob centroid is deleted from the list.

V. ALGORITHM

- Maximum number of objects we will be tracking are 200.
- System object for reading the video with each frame of the type 'Single'. For this purpose, we have used vision.VideoFileReader function.
- Offsets for drawing bounding boxes in original input video. Here, we have bounded all moving objects in boxes, for this we used int32(repmat([roi(1), roi(2), 0, 0],[maxNumObj 1]));
- In the next step, RGB image is converted into YCbCr using vision.ColorSpaceConverter('Conversion', 'RGB to YCbCr');
- Next step is background subtraction.
- Remove noise and really small blobs using vision.MorphologicalClose('Neighborhood', strel('square',10));
- Find the blobs in the segmented images, properties of blobs also specified.
- Create system objects for players with their locations.
- Will run a loop till all the frames are completed, storing all the frames and background.
- In this step, background is subtracted on the luminosity, Chroma part of YCbCr.
- Fill in small gaps in the detected objects and club the separated image.
- Blob analysis, storing area, centroid, bounding box in a list.
- Traversing through the list of centroids of all the blobs.
- Will check if the position of a particular blob has changed or not, on the basis of the area of blob, and distance between the centroid of that blob in the background frame and the current frame, if yes, increase its count by 1, if not, increase its miss count by 1.
- If the stationary object remains there for 75 frames, it will be declared as an abandoned object, while if the stationary object, moves or occluded by some interference for more than 7 frames, it will be no longer stationary object.

- As soon as the object becomes an abandoned object, its bounding box becomes red, and an alarm is raised.

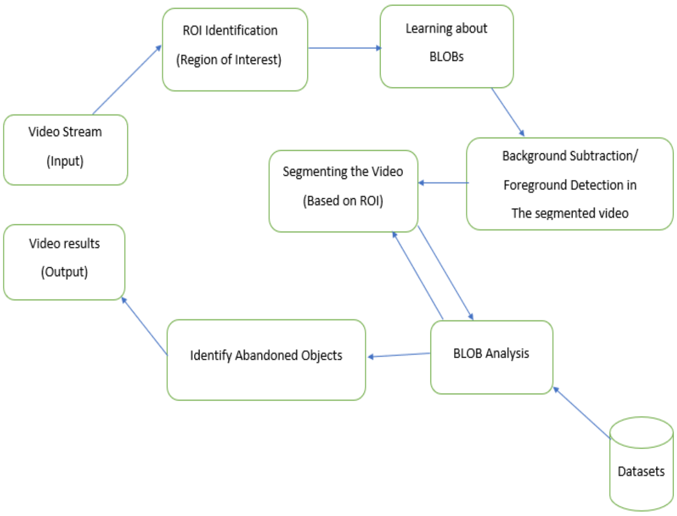


Fig 1:- Algorithmic working sequence of the project

VI. RESULT

The Proposed algorithm in previous section has been implemented as real-time system. The algorithm was performed on standard database named ABODA(Abandoned Object Dataset) .

MATLAB R2016a was used for video processing and analysis. No specific extra hardware wasn't used. HP au111tx, with Intel i5 7th generation core processor was enough to handle the heavy MATLAB application.

The algorithm was tested on the dataset and the results were recorded as follows.

Fig 2. Abandoned Object Detected

Fig 3:- YCbCr form of the aboe image

Fig 4:- Binary Image ater image subtraction from the background(Without removing noise)

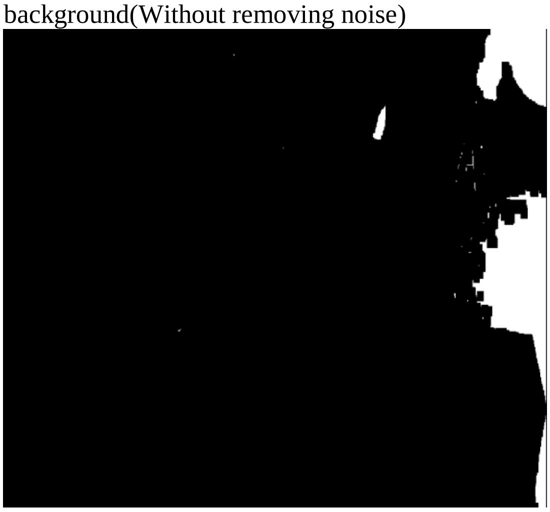


Fig 5:-Binary image obtained after removing the noise from Fig 3

TABLE I. TEST ON DATASET

Table Column Head		
Data Set No.	***Abandoned Object Present***	***Abandoned Object Detected***
Video1	Yes	Yes
Video2	Yes	Yes
Video3	Yes	Yes
Video4	Yes	Yes
Video5	Yes	Yes
Video6	Yes	No
Video7	Yes	No
Video8	Yes	No
Video9	Yes	Yes
Video10	Yes	Yes

Fig 6:-Results of al the videos of dataset

VII. CONCLUSION

This paper proposed a system for automatic Abandoned object detection without any need for human manpower for monitoring the system. Such system proves to be efficient in public place for providing security. The application was tested on an official as well as an unofficial dataset and was found to give accuracy comparable to most contemporary systems. One particularly noteworthy aspect of its performance was that it was able to detect the object in real time aspect with video speed almost equal to the real-time aspect.

VIII. LIMITATIONS

- If a person doesn't moves for a long period of time then it can also be declared as Abandoned Object.
- The project is inefficient for lightning change. It may mistake some shadows as an abandoned object.
- The system doesn't checks the proximity of the person from the object. Thus if a person stands along with that object kept on floor, the system will declare it as abandoned object.
- It won't work efficiently for extremely poorly lit videos.
- Background is not updated.
- Since only the first frame is seen as the background , so if a person or object is there in the first frame and then it is removed ,it is treated as an abandoned object.

IX. FUTURE WORK

Learning rate can be increased so that it can figure out the complete background using the first few frames.
Also it can help to adapt to lighting changes.
In future we can go for object classification to avoid false detection.
Machine Learning feature can be added so as to prevent the detection of a stationary person as an abandoned object.
Further, detection using proximity can also be used before declaring an object as abandoned. That is, until the person who owns the object is within a particular radius the object would not be declared as abandoned object.
Multiple numbers of cameras can be used for more efficient analysis of the particular area.

ACKNOWLEDGMENT

I would like to thank Prof. Dr. Anish Chand Turlapaty and T.A. Viswanatha Reddy who gave me such a great opportunity to work on this project. I am also thankful to my colleagues, Paras Dahiya, Pradeep Turan, Sarin Nikose, who supported me on each and every moment to out pass all the hurdles in the implementation and execution of this project.

REFERENCES

[1] Aditya Gupta, Neeraj Dhanraj Bokde, Vishal Satpute, K.D. Kulat --- Real Time Abandoned Object Detection Using Video Surveillance published in January 2015

[2] Abhineet Kumar Singh –B.Tech IIIT Allahabad final year project report 2009 under the professor Dr. Anoopam Aggarwal

[3] C. Stauffer and W. E. L. Grimson, ―Adaptive background mixture models for real-time tracking,‖ in Proc. IEEE Conference on Computer Vision and Pattern Recognition, Vol. 2, pp. 246–252, Feb. 1999. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[4] Y. Tian, M. Lu and A. Hapapur, ―Robust and efficient foreground analysis for real-time video surveillance,‖ in Proc. IEEE Conference on Computer Vision and Pattern Recognition, June 2005.

[5] A.Singh, S.Sawan, M.Hanmandlu, V.K.Madasu, B.C.Lovell ―An Abandoned Object Detection System Based On Dual Background Segmentation,‖ in Proc. Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, 2009.

[6] R. Singh, S. Vishwakarma, A. Agrawal and M. D. Tiwari, ―Unusual activity detection for video surveillance,‖ in Proc. 1st International Conference on Intelligent Interactive Technologies and Multimedia, Dec. 2010.