

Abandoned Object Detection in Educational Institutes using Video Surveillance

Avinash A. Utikar¹, Dr. Nilesh J. Uke²

Research Scholar, Department of Information Technology, Sinhgad College of Engineering, Pune, India¹

Professor, Department of Information Technology, Sinhgad College of Engineering, Pune, India²

Abstract: Surveillance cameras are video cameras used for the purpose of observing an area. Sensitive areas need to be observed continuously for security reasons. The security of these areas shall not be depending on the size of an area but, shall be considered on the high priority of security. Before few decades this task was done by the humans, but now it is not possible for human to keep watch day and night (24x7) on such areas. We all know that there is massive increase in theft in either in one or more ways. Human can easily be distracted and, a small distraction is sufficient for thieves to break the security. System is proposed and presented using our own data set captured in laboratory datasets. The proposed system works on QVGA (Quarter Video Graphics Array) resolution at which most CCTV (closed circuit television) cameras operate and uses a simple mathematical model for detection of human and object. Proposed work is a combination of base works as 1: Real time video feed. 2: Discover frame from the video after fixed time interval. 3: Dual time background subtraction algorithm which periodically updates two sets of background. 4: Deriving a reference image. 5: Maintaining a Buffered Background image/frame. 6: Blob detection. 7: Motion Tracking. 8: Object tracking. 9: Detection and maintaining log of and Alarm signalling. An algorithm for tracking of abandoned objects even under occlusion is also proposed.

Keywords: Video Surveillance System, Image Separation, Noise Removal, Human Detection, left baggage detection, background segmentation, tracking.

I. INTRODUCTION

Video surveillance system for maintaining the security at all possible level is being used by human since last two decades. In early phases of video surveillance system the video captured by the cameras (on site) would simply be displayed on the screen present in the control room. Continuous monitoring of displayed video is done by human recourse for any abnormal activities. In recent years it has been observed that there is massive increase in terrorist attacks on public places. Train stations, bus stations, airport, subways, shopping malls, markets are the most targeted areas of these attacks. Even the CCTV surveillance systems are present in these areas, but the analysis of footages that has been captured by these CCTV's, are rarely observed by humans. Recent studies explore and analyze the limitations or effectiveness of human being in analyzing and processing the footages particularly of crowded scenes in surveillance system consisting of many cameras[1-3].

Automatic video surveillance is becoming increasingly important in many applications, including traffic control, urban surveillance, home security and healthcare. With the innovation of intelligent (smart) cameras having greater processing capabilities it is possible to design the system that can detect the suspicious behaviors and abandoned objects. There are many algorithms suggested by many researchers to handle the problem human tracking and abandoned object detection. In practice many of these algorithms are based on complex mathematical manipulation to solve simple problems, because of this many of these algorithms failed to perform adequately. This work presents a system based on a simplistic and instinctive mathematical model that works efficiently with

standard input which is the industry standard for most CCTV cameras. The proposed work consists of a dual background subtraction technique which is based on Approximate Median model framework. Algorithms for tracking abandoned objects with or without occlusion are also proposed.

II. LITERATURE SURVEY

An Basically there are three conventional approaches to moving object detection: temporal differencing, optical flow and background estimation methods. Temporal differencing [4] is a technique which work better for environments that are dynamic in nature, but the same technique is not so good when need to extract the relevant quality pixels[6].The technique used to detect an independently moving object in presence of motion camera is Optical flow method [4], [5], [7]. In practice the optical flow methods are complex and not applicable for real-time algorithm.

In the next part of system few algorithms and methodologies have been proposed for Abandoned Object Detection. A. Singh has proposed the some method on dual background segmentation in which blob detection, tracking is done [1].A robust method has been proposed for recognizing abandoned object detection and also proposed a method based on illumination of foreground mask[8]. Objects that are deserted and removed from the site need to be handled; the solution for this situation is evolved by Lin et.al [8]. For real-time environment two methods are proposed based on Gaussian mixture model for object detection among two one is efficient and another one is robust [2].

A. Page Layout

Your paper must use a page size corresponding to A4 which is 210mm (8.27") wide and 297mm (11.69") long. The margins must be set as follows:

- Top = 19mm (0.75")
- Bottom = 43mm (1.69")
- Left = Right = 14.32mm (0.56")

Your paper must be in two column format with a space of 4.22mm (0.17") between columns.

III. SYSTEM OVERVIEW

A. System Design Overview

Overall design of the proposed system is as shown figure 1. It consists of few modules connected, that describe the operational flow of the system. Every module is designed in modular method, it helps to easily modify respective module. This method will help to maintain the compatibility of input and output data types used with respective module. The modules are separated as in processing blocks. The blocks are: Data extraction and conversion unit; Background subtraction module; object tracking and occlusion detection block and Alarm rising and display of result unit.

A live video stream is provided as an input to the system, this video stream is segmented in to images or we can say that a sequence of images. With constant time interval the sequence of images is mapped in to a frame. These frames help to discover the reference image for initial reference background. Now considering the pixel features height, width and intensity of each pixel a 3D intensity matrix (for reference image) is passed as an input to the Background Subtraction module.

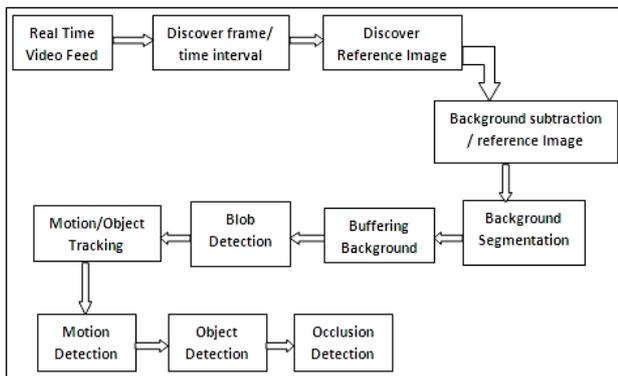


Figure 1: Overall system operation

B. Background Segmentation

There are many authors who have suggested many background subtraction methods. Based on the Gaussian mixture model is most popular method which was suggested by Friedman and Russell [9] first time. Several other authors modified this method according to the need of their work. An algorithm is proposed based on Approximate Median algorithm for background subtraction. This method is dynamic, easily adaptive, and instinctive in nature. In this work the information about pixel like color/intensity is used for background processing. As this system maintain the two different background frames, so the complex mathematical

manipulation is reduced and, this will help to maintain system self-adaptability for real-time.

In this work the moving object or crowd, changes in light effects etc. are considered for maintaining the buffered background and current background with specific time interval for motion detection and object detection.

C. Algorithm

Based on Approximate median model for separation of foreground and background an algorithm is proposed in this system [10]. Two reference background images are considered for comparison purpose. These two are named as Buffered background and Current background.

This background separation technique is called as a dual background separation method.

First background is initially stored and updated to buffered background and the second background is updated dynamically and updated to current background image.

The first frame of the incoming video is referred as 'Current Background'. Consequently, the intensity of each pixel of this current background is compared with the corresponding pixel of the next frame (after every 0.4 seconds). If it is less, then the intensity of that pixel of current background is incremented by one unit, otherwise it is decremented by one unit. In case of equality, the pixel intensities remain unchanged. This way, even if the foreground is changing at a fast pace, it will not affect the background but if the foreground is stationary, it gradually merges into the background.

Stored set of background images called as 'Buffered Background'. Here, all those pixels which do not belong to the prospective objects set are made equal to that of 'Current Background'. This is done at an interval of every 20 seconds. Difference of the two backgrounds is represented as a binary image with the white portion representing foreground (blobs).

IV. OBJECT DETECTION

In this module, the binary image divide from the previous unit into a number of legitimate blobs (rectangular regions enclosing continuous regions of foreground). Once the blobs and their various properties like area, centroid position etc. have been generated, the tracking algorithm is applied.

A. Mathematical Model

Let us suppose that, after blob analysis, we get 'N' number of blobs, each with its enclosing region $R_n(t, l, h, w)$, its area ' A_n ' and centered $C_n(i, j)$.

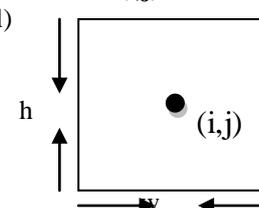


Figure 2: A typical blob

Where, t is the pixel value at top position; l is the pixel value at left position; h is the blob height; and w is the blob width; and $1 \leq n \leq N$

Let 'T' be the set of tracked blobs such that,

$$T = [B_n : B_n = \{ R_n(t, l, h, w), A_n, C_n(i, j), tn, m_n \}]$$

$$\forall 1 \leq n \leq M$$

Where,

M is the tracked blobs number; **tn** is the frames for which the blob has been tracked and **m_n** is the successive frames for which the blob being tracked.

After analysis of the present frame we got a set of blobs let us represent it with 'P', and its objects as **bn**, which are **N** in number. Then the set of blobs is

$$T = [b_n : b_n = \{ r_n(t, l, h, w), a_n, c_n(i, j), tn, m_n \}]$$

$$\forall 1 \leq n \leq N$$

B. Blob detection

Binary image is provided as an input to an algorithm one which described and returns various properties of the detected blobs as bounding box, area, centred place etc. A simplified version of the algorithm is as follows:

1. Discover and create a region counter.
2. An image is scanned from all four directions, from left to right and from top to bottom.
3. Check for every pixel the north and west pixel (4-connectivity) or the northeast, north, northwest, and west pixel (8-connectivity) to find out intensity value of 1 in the binary image (termed as criterion of blob analysis).
4. If the neighboring pixel intensity is not fitting into the criterion assign a value to region counter. Increment the region counter.
5. When any one neighboring pixel fits into the criterion, that pixel is assigned to that region.
6. With multiple regions, if more than one neighbor matches the criterion of the same region then assign pixel to their respective region.
7. If the members from different regions match the criterion, assign those to any one of region amongst present and notify those of that region and are equivalent.
8. Image is scanned again and the same value is assigned to all the regions those are equivalent.

C. Tracking

The next process in object detection is tracking the different blobs so as to find which blobs correspond to abandoned objects. The first step in this process is to create a set, Track, whose elements have three variables: blob- Properties hit Count and miss Count. The next step is to analyse the incoming image for all the blobs. If the area change and the centroid position change, as compared to any of the elements of the set Track are below a threshold value, increment hit Count and reinitialize miss Count with a zero; otherwise tried to create a new element in the Track-set, initializing the blob-properties variable with the properties of incoming blob and hit Count and miss Count are initialized to zero. Then run a loop through all the elements of the set. If the hit Count goes above a user defined threshold value, an alarm is triggered. If the

miss Count goes above a threshold, delete the element from the set. These two steps are repeated until there are no incoming images.

D. Pseudo Code for Tracking

Take area, centroid, bounding boxes (bbox) and the total number of blobs (n) as input from Blob Analysis block. Let Track=empty set of vectors of type t where t=(area, centroid, bbox, hit count, miss count, active, occluded)

```

m= Track.Size
For i=1 to n
    c=0
    For j=1 to m
        If (percentage background in Track[j].area<50)
            Then Track[j].occluded=true
        End
        If (|area[i]-Track[j].area/area[i] <.05 and
|centroid[i]-
Track[j].centroid/centroid[i]<.05)
            Then Track[j].active =true, c=1,
            break from loop
        End
    End
End
If c=0
    Then k=Track.size++,
    Track[k].area = area[i]; Track[k].centroid = centroid[i];
    Track[k].bbox = bbox[i]; Track[k].hitcount = 1;
    Track[k]. active = true;
End
End
m= Track.Size
For j=1 to m
    If (Track[j].active==true)
        Then
            Track[j].hitcount=Track[j].hitcount+1;
            Track[j].misscount=0;
            If (Track[j].hitcount > 4)
                Don't update pixels of Track[j].bbox in buffered
                background
            End
            If (Track[j].hitcount > 40)
                Then raise alarm for Track[j]
            End
            If (Track[j].active==false and misscount >3)
                Then delete Track[j]
            End
        End
    End
End
Update the buffered background

```

E. Occlusion Detection and Tracking

A tracked blob is considered to be occluded if its major region (say 80 %) is covered by foreground and it should continue to be tracked if either it is occluded or its area and centroid is matched with any of the blobs of set P. An alarm is raised if **tn** > threshold. The blob is removed from T if **mn**>3. This idea is similar to the method used in and for occlusion detection, but instead of keeping track of two different foregrounds, system proposes the following modification.

Let us assume that a particular portion of the frame containing the blob which is being tracked (i.e. present in the 'Track' set of blobs) is now occluded. Due to this occlusion, the blob signifying that particular object won't be included in the present set of Track.

Following are the possibilities in the new frame of the blob that was being tracked up to previous frame: -

- a) Object is removed from the location. In this case, the blob area representing the object should contain background pixels.
- b) There may be a new object at the same location.
- c) There is a new object which completely or partially occludes the old object.

An exception to the above cases is when a tracked object is removed while being occluded or another object of similar size is placed in camera's line of view. To deal with occlusion we have added the following two steps to the tracking algorithm:

Step 1: Calculate the number of pixels of buffered background which are same as that of current background for that element of set Track which has suddenly stopped being tracked (due to occlusion, or removal from scene). If it's below a threshold value, say 50 percent, and the hit Count is above a threshold value (making sure the blob has been tracked long enough), we label this element of the set Track as occluded.

Step 2: Go to step 2 of the tracking algorithm. A blob labeled occluded remains in the Track set; i.e. its hit Count is incremented and missCount is reinitialized. Rest of the tracking algorithm remains same.

F. Alarm and Display

System use the Raise-alarm flag from previous units and highlight that part of the video for which the alarm has been raised. System also display the binary-image (without background) video so that the operator can fine tune the value of for shadow and reflection subtraction.

V. CONCLUSION

This Proposed work helps for security of educational institute under video surveillance, system based on a dual background segmentation scheme. Blob analysis is done on the segmented background and a dynamic tracking algorithm is devised for tracking the blobs. Proposed work is simple and computationally less intensive as it avoids the use of expensive filters while achieving better detection results. Still the research is going on in the detection techniques.

REFERENCES

- [1] A. Singh, S. Sawan, M. Hanmandlu, V.K. Madasu, B.C. Lovell "An abandoned object detection system based on dual background segmentation" IEEE 2009 Advanced Video and Signal Based Surveillance.
- [2] Xuli Li, Chao Zhang, Duo Zhang, "Abandoned Objects Detection Using Double Illumination Invariant Foreground Masks", 2010 International Conference on Pattern Recognition, 1051-4651/10, 2010 IEEE.

- [3] Murat EKINCI, Eyup GEDIKLI, —Silhouette Based Human Motion Detection and Analysis for Real-Time Automated Video Surveillance||, VOL.13, NO.2 2005, c TUBITAK, pp. 199-229
- [4] Nan Lu, Jihong Wang, Q.H. Wu and Li Yang, —An improved Motion Detection Method for Real-Time Surveillance||, IAENG International Journal of Computer Science, 35:1, IJCS_35_1_16
- [5] Denman, Simon P. and Chandran, Vinod and Sridharan, Sridha (2007) An adaptive optical flow technique for person tracking systems. Pattern Recognition Letters 28(10):pp. 1232-1239
- [6] Peng Zhang, Tie-Yong Cao, Tao Zhu, —A Novel Hybrid Motion Detection Algorithm Based On Dynamic Thresholding Segmentation||, Communication Technology (ICCT), 2010 12th IEEE International Conference, pg: 853-856
- [7] Yokoyama, M.; Poggio, T, —A contour-based moving object detection and tracking||, IEEE 2005, pp. 271-276
- [8] Chih-Yang Lin and Wen-Hao Wang, "An Abandoned Objects Management System Based on the Gaussian Mixture Model", International Conference on Convergence and Hybrid Information Technology 2008 978-0-7695-3328-5/08, 2008 IEEE.
- [9] N. Friedman and S. Russell, "Image segmentation in video sequences: a probabilistic approach", in Proceedings of 13th Annual Conference on Uncertainty in Artificial Intelligence, 1997, pp. 175-181.
- [10] N.J.B. McFarlane and C.P. Schofield, "Segmentation and tracking of piglets in images", Machine Vision and Applications, vol. 8, 1995, pp. 187-193.