Abandoned Object Detection with Video Surveillance

by

CHOWDHURY FARSAD AURANGZEB - 11301019

ISRAT JAHAN SHANCHARY - 10201022

SALMAN TASDID - 08110072

Supervisor

RUBEL BISWAS

Statement

We hereby proclaim that this thesis is based on the results we found by our hard work. Contents of the work found by other researcher(s) are motioned by references. This thesis has never been previously submitted for any degree neither in whole nor in part.

Signature of the Author(s):

____________________	____________________	____________________
Chowdhury Farsad Aurangzeb	**Israt Jahan Shanchary**	**Salman Tasdid**
11301019	**10201022**	**08110072**

Signature of thesis Supervisor:

____

Rubel Biswas

Acknowledgement

At first we want to thank the school of computer science and Engineering for providing us such opportunity to express our knowledge and learning of all this four years of student life. We would also like to thank all the teachers of Computer science and engineering department for their direct and indirect support. It would be unfair no to thank all our friends and fallow mates for their extreme encouragement and suggestions from their previous experience. Our special and heartily thanks to our research supervisor Rubel Biswas and co-advisor Md. Zahangir Alom for the way they directed us and brought this work towards a completion. It is with their supervision that this work came into existence.

Abstract

We are presenting a new framework for smart threat detection with the help of video surveillance which captures live streaming from the camera and determine if any abandoned object has been set up. It then raises an alarm right after the carrier of the bag leaves the scene. The abandoned object is detected with the help of background subtraction and brought more accuracy on the object's image with the help of morphology. Every individual person on the frame is being detected by their body with the help of Kalman filter and brought relation of the body with the bag, so that the alarm is raised only when the body of the abandoned object carrier leaves the frame.

Table of Content

List of Figures

1. Introduction

1.1 Introduction

Surveillance system has become an important aspect in security and a necessity to keep proper check. There was a time when video surveillance was mostly used by the government and big companies. But now the scenario has changed. Presently the use of surveillance camera has increased and it is increasing more. Everywhere you look, whether it’s a restaurant, office, educational institution or station, presence of video surveillance is almost obvious. This is because of technology being much more inexpensive and the awareness of people to take precautions to a crime than a cure. But the criminals are as aware as everyone else. So, to guard from any unpredicted situations from occurring we need the surveillance system to be more advanced and accurate. That is how the idea of abandoned object detection through video surveillance came up. With the help of which we can detect if any suspicious package is left by anyone and then raises an alarm. It is challenging to detect an object when the color of the object almost matches the color of the background. Here, the object will still be successfully detected in our project.

1.2 Motivation

Security and video surveillance has turned into a growing market. There will always be thieves and there will always be criminals. That is why there will always be a huge demand of new security cameras with advanced security technology to compete against the ever evolving ways of crime that people come up with. A breathtaking example is the pressure cooker bomb’s explosion in Boston Marathon in April 15, 2013 by Chechen brothers Dzhokhar Tsarnaev, Tamerlan Tsarnaev though the Officials swept the area for bombs twice before the explosions; the second sweep occurred one hour before the bombs went off. People were able to come and go freely, and carry bags and items in and out of the area, which is why the Chechen brothers got in with their bag pack and abandoned it which could not be detected by neither the surveillance camera nor the eyes of the security personals. In such circumstances abandoned object detection would be a great defender. The saying, “An ounce of prevention is worth a pound of cure” applies here. When dealing with crime the goal is always prevention, and having a video surveillance system can certainly help. Nobody can deny the endless applications that come with the use of security cameras and surveillance technology. That is what triggered our thought regarding working on a project that helps alarming people before any unexpected situation takes place. It is all in attempts to create a safer environment, lowering crime, and most importantly keeping your valuable life, business and assets protected.

1.3 Related Work

Physical security being one of the top priorities, advanced surveillance system has recently attracted considerable interests. Surveillance system identifying abandoned object is considered to be a great contribution for security and safekeeping. But even today most of the surveillance company provides surveillance camera along with a person who will sit at the back end and see through the video footage whether any suspicious activities occurring or not. This process is no way optimal. Several models have been proposed based on a variety of techniques along with some limitations with each.

1.4 Research Methodology

This research is basically aimed for implementing a smart surveillance system that is capable of detecting any abandoned object left by someone. It also identifies the owner by detecting the body. When the body is found to be missing in the region of interest, an automatic alarm is raised to aware the authority. The whole thesis is done in Matlab 2014a where the object detection is done by Background subtraction and the body detection is done by Kalman filter. The image of the object is brought into more accuracy with the help of Morphological Closing. To verify the completion of our project we made several videos where we made people leave a bag in the middle of the video frame. Most of the videos were made including multiple people and all videos were made in different lighting condition to evaluate the role of lighting in the accuracy of detecting the abandoned object and its carrier. We used mobile camera and Digital Single Reflex Camera (DSLR) to make these videos and used video editing software to cut short and edit these videos.

1.5 Outline

This report is presented in five chapters. Chapter 1 contains introduction, motivation, related works and research methodology. Chapter 2 attempts to represent the proposed approach which contains Background subtraction, Threshold and blob analysis, Region of interest, Morphology and Kalman filter. Chapter 3 contains Results and discussion where the accuracy of our work is shown by different graph and table. Chapter 4 contains conclusion and future work. The report is ended in chapter 5 with reference web links and published article

2. System Implementation

2.1 Problem Definition

Real time Video has been captured using Walton Primo 5 mega pixel mobile camera and Canon 1100D DSLR (12 mega pixel) and it has been processed into frames. These frames have been used as input images for the system. The goal is to track an abandoned object in a crowded place. Here, we have used Background Subtraction algorithm to track the abandoned object where we are recording the initial frame as the background. We have also assigned a region of interest (ROI) where the possibility of object abandonment is high. As the surveillance is for security purpose and accuracy is most necessary for these reason, we have used Morphology to enhance the accuracy for tracking the abandoned object. We also used Kalman filter to track the persons entering into the frame and leaving and relate them with the abandoned object.

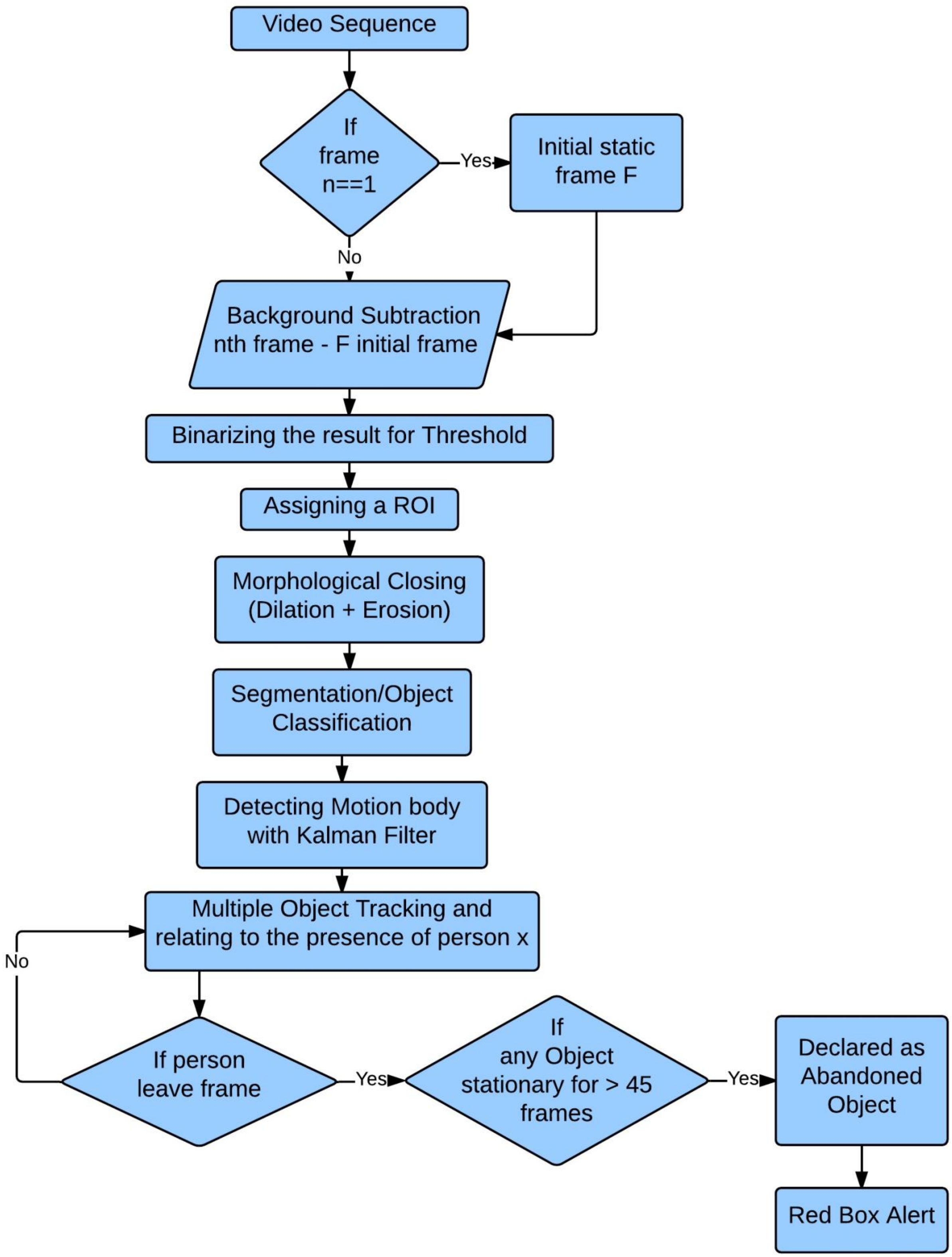


Figure 1: Flow chart of our Video Surveillance System

2.2 Proposed Approach

I. Object Detection with Background Subtraction.

II. Assigning Region of Interest (ROI).

III. Using Morphology to eliminate gaps between pixels of any detected object for more accuracy and smoothens of the object.

IV. Using Kalman Filter to detect persons in the frame and relate them with the abandoned object.

2.2.1 Background Subtraction

In the field of object detection, background subtraction plays a vital role. There are lots of proposed algorithms regarding background subtraction, for example, Mean filter, Gaussian Average, Background Mixture model. Background subtraction (BS) is a common and widely used technique for generating a foreground mask (namely, a binary image containing the pixels belonging to moving objects in the scene) by using static cameras. As the name suggests, BS calculates the foreground mask performing a subtraction between the current frame and a background model, containing the static part of the scene or, more in general, everything that can be considered as background given the characteristics of the observed scene.

Here, first we have worked for calculating the image containing only the background, a series of preceding images are averaged. For calculating the background image at the instant *t*,

$$B(x,y) = \frac{1}{N}\sum_{i=1}^{N} V(x,y,t-i)$$

where N is the number of preceding images taken for averaging. This averaging refers to averaging corresponding pixels in the given images. N would depend on the video speed (number of images per second in the video) and the amount of movement in the video.1111 After calculating the background $B(x,y)$ we can then subtract it from the image $V(x,y,t)$ at time t=t and threshold it. Thus the foreground is

$$|V(x,y,t) - B(x,y)| > \mathrm{Th}$$

where Th is threshold. Similarly we can also use median instead of mean in the above calculation of $B(x,y)$.

Here, we distributed our Background modeling into two main steps: Background Initialization and Background Update. In the first step, an initial model of the background is computed, while in the second step that model is updated in order to adapt to possible changes in the scene. We now classified a pixel as background if its current intensity lies within some confidence interval of its distribution's mean:

$$\frac{|(I_t - \mu_t)|}{\sigma_t} > k \longrightarrow Foreground$$

$$\frac{|(I_t - \mu_t)|}{\sigma_t} \leq k \longrightarrow Background$$

where the parameter k is a free threshold (usually $k = 2.5$). A larger value for k allows for more dynamic background, while a smaller k increases the probability of a transition from background to foreground due to more subtle changes[16].

Here, our motion detection algorithm starts with the segmentation part where foreground or moving objects are segmented from the background. We implemented this by taking the initial frame as the background and take the frames obtained at the time t, denoted by I(t) to compare with the background image denoted by B. Here using simple arithmetic calculations, we have segmented out the objects simply by using image subtraction technique of computer vision meaning for each pixels in I(t), then taking the pixel value denoted by P[I(t)] and subtracting it with the corresponding pixels at the same position on the background image denoted as P[B].

In mathematical equation, it is written as:

$$P[F(t)] = P[I(t)] - P[B]$$

The background is assumed to be the frame at time *t*. This difference image would only show some intensity for the pixel locations which have changed in the two frames. Though we have seemingly removed the background, this approach will only work for cases where all foreground pixels are moving and all background pixels are static.

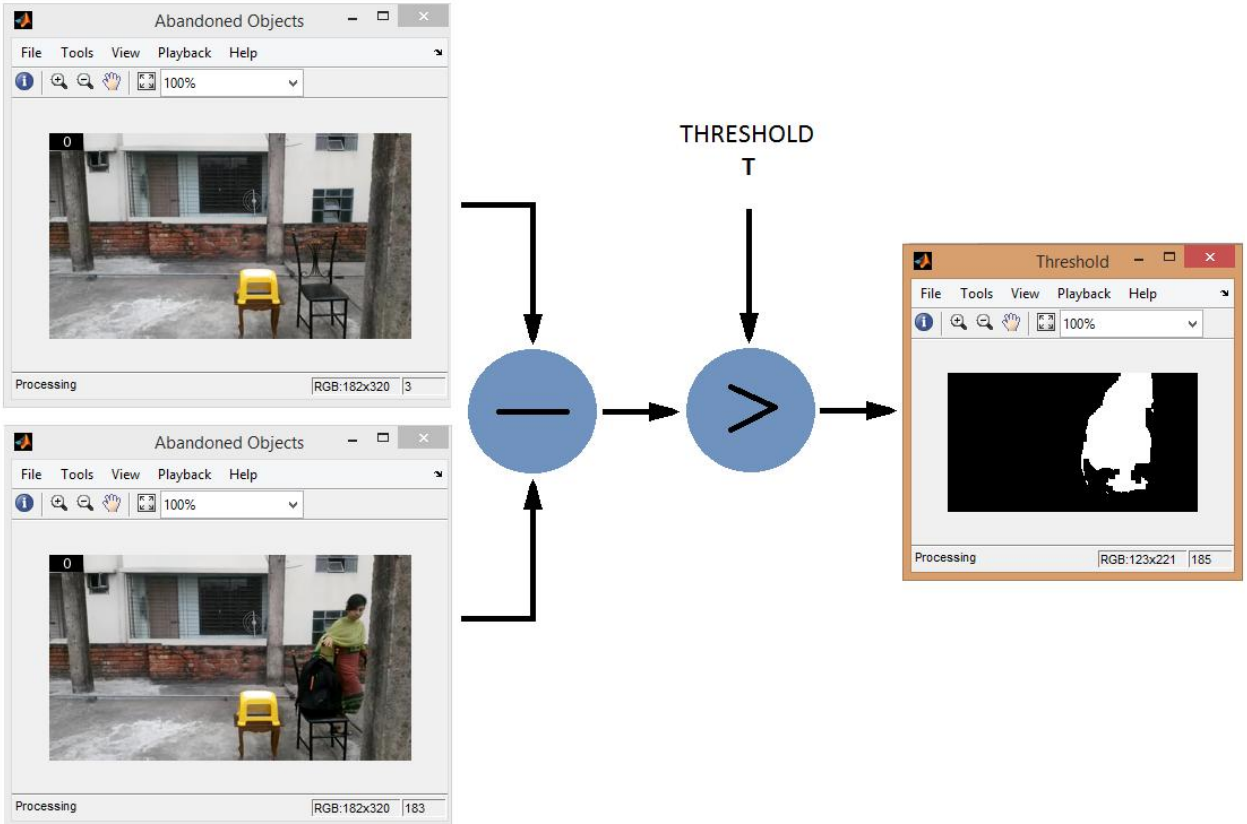


Figure 2: How Background Subtraction is working in our system

2.2.2 Threshold and Blob Analysis

A threshold "Threshold" is put on this difference image to improve the subtraction.

$$|P[F(t)] - P[F(t+1)]| > \text{Threshold}$$

This means that the difference image's pixels' intensities are 'thresholded' or filtered on the basis of value of Threshold[16].

Setting a desired threshold value as a standard and converting a colored image into values of 0 and 1 is called “binarization processing” (usually black is 0 and white is 1). The method of analyzing an image which has undergone binarization processing is called “blob analysis”. A blob refers to a lump. Blob analysis is image processing's most basic method for analyzing the shape features of an object, such as the presence, number, area, position, length, and direction of lumps. Inside this subsystem, the Luminance Segmentation and Color Segmentation subsystems perform background subtraction using the intensity and color data. Here, we combined these two segmentation results using a binary OR operator. The Blob Analysis block computes statistics of the objects present in the scene. By using this process it is possible to focus on only the object of inspection, and conduct various analyses.

Once we obtain the region that corresponds to the object or the objects being inspected, we may commence the analysis - that is, extract the information we are interested in. Our program right now is capable of extracting the region that directly corresponds to the capsules visible in the image. What remains is to inspect each capsule and classify it as a correct or deformed one.

As we want to analyze each capsule separately, we have start with decomposition of the extracted region into an array of connected components (blobs). This common operation can be performed using the straightforward SplitRegionIntoBlobs filter.

2.2.3 Region of Interest (ROI)

Let us consider a video sequence where there is both a luggage counter and a potential place where an object can be abandoned by a potential threat. Here our surveillance system should focus on only the region where the possibility of an object being abandoned is high, not the luggage counters where our system will continuously detect bags. So, to overcome this problem we have extracted a ROI (Region of Interest) for which the video areas are eliminated which are most unlikely to contain abandoned object. The ROI defines the borders of an object under consideration. In many applications, symbolic (textual) labels are added to a ROI, to describe its content in a compact manner. In our thesis it is done by generating a particular code resulting the region of interest. A *region of interest* (ROI) is a portion of an image that one would want to filter or perform some other operation on. ROI is defined by creating a *binary mask*, which is a binary image that is the same size as the image to be processed with pixels that define the ROI set to 1 and all other pixels set to 0.

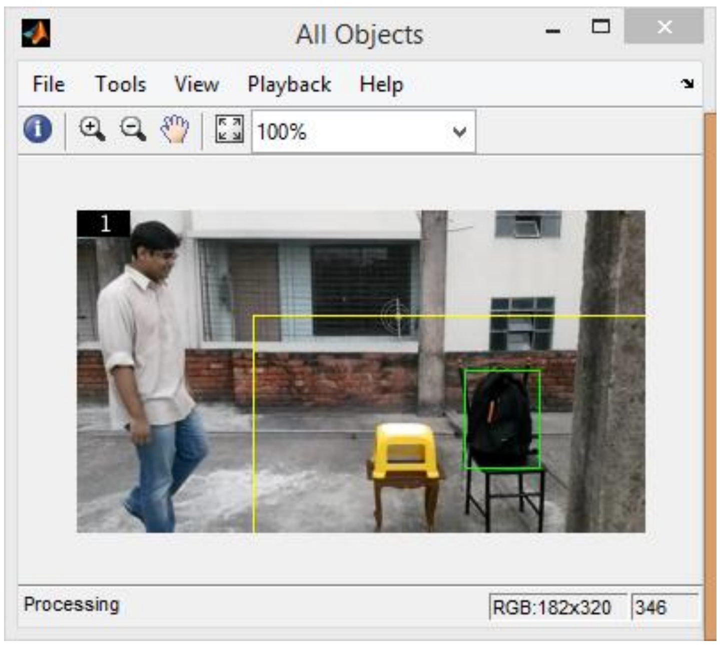


Figure 3: Region of Interest

This is the image of our Project that shows all Objects window marks the region of interest (ROI) with a yellow box and all detected objects with green.

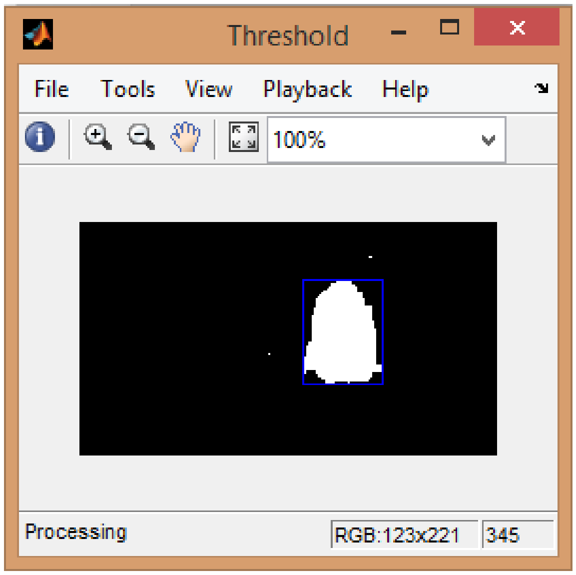


Figure 4: Result of BS in ROI with blue box

The Threshold window shows the result of the background subtraction in the ROI with a blue box.

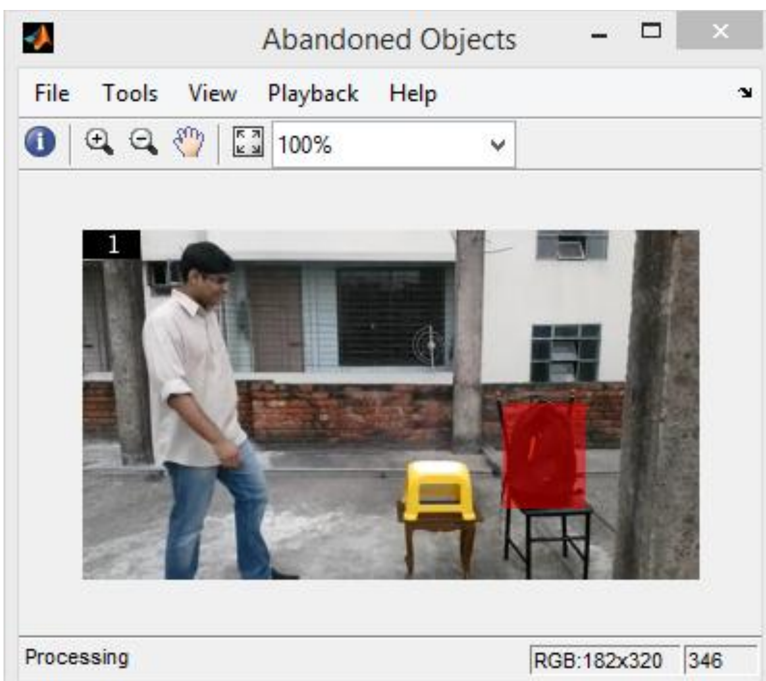


Figure 5: The Abandoned Objects window highlights the abandoned objects with a red box.

2.2.4 Morphology

Now we can see that if the color of the abandoned object and the background is same in that case the object is detected in a very unclear way. In such cases the pixels break and a single object may be detected as several objects. Now before describing what we proposed to eliminate this problem, we are going to explain what Morphology is.

Morphological image processing is a collection of non-linear operations related to the shape or morphology of features in an image. Morphological operations rely only on the relative ordering of pixel values, not on their numerical values, and therefore are especially suited to the processing of binary images. Morphological operations can also be applied to greyscale images such that their light transfer functions are unknown and therefore their absolute pixel values are of no or minor interest.

Morphological techniques probe an image with a small shape or template called a **structuring element**. The structuring element is positioned at all possible locations in the image and it is compared with the corresponding neighbourhood of pixels. Some operations test whether the element "fits" within the neighborhood, while others test whether it "hits" or intersects the neighborhood:

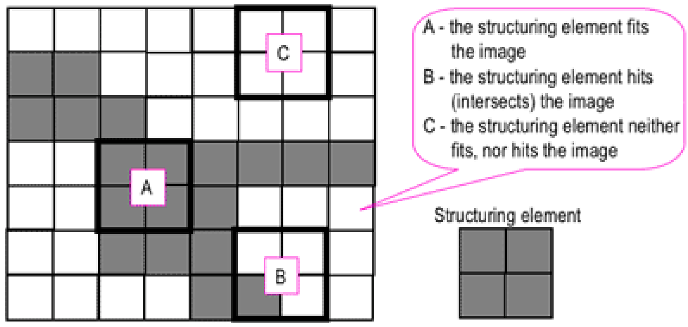


Figure 6: Probing of an image with a structuring element

(white and grey pixels have zero and non-zero values, respectively).

A morphological operation on a binary image creates a new binary image in which the pixel has a non-zero value only if the test is successful at that location in the input image.

The **structuring element** is a small binary image, i.e. a small matrix of pixels, each with a value of zero or one:

- The matrix dimensions specify the *size* of the structuring element.
- The pattern of ones and zeros specifies the *shape* of the structuring element.
- An *origin* of the structuring element is usually one of its pixels, although generally the origin can be outside the structuring element.

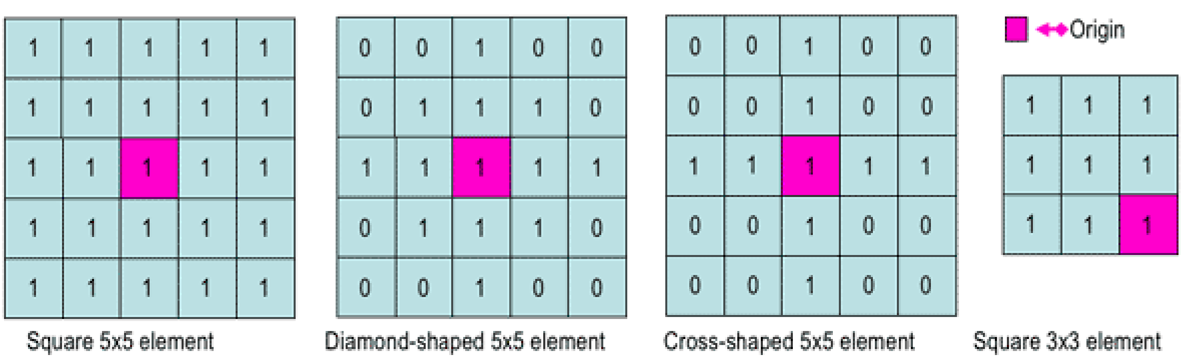
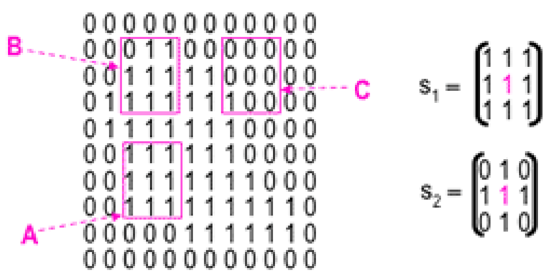


Figure 7: Examples of simple structuring elements.

A common practice is to have odd dimensions of the structuring matrix and the origin defined as the centre of the matrix. Stucturing elements play in moprphological image processing the same role as convolution kernels in linear image filtering.

When a structuring element is placed in a binary image, each of its pixels is associated with the corresponding pixel of the neighbourhood under the structuring element. The structuring element is said to **fit** the image if, for each of its pixels set to 1, the corresponding image pixel is also 1. Similarly, a structuring element is said to **hit**, or intersect, an image if, at least for one of its pixels set to 1 the corresponding image pixel is also 1.



		A	B	C
fit	s_1	yes	no	no
	s_2	yes	yes	no
hit	s_1	yes	yes	yes
	s_2	yes	yes	no

Figure 8: Fitting and hitting of a binary image with structuring elements s_1 and s_2.

Zero-valued pixels of the structuring element are ignored, i.e. indicate points where the corresponding image value is irrelevant.

Fundamental operations

More formal descriptions and examples of how basic morphological operations work are given in the Hypermedia Image Processing Reference (HIPR) developed by Dr. R. Fisher et al. at the Department of Artificial Intelligence in the University of Edinburgh, Scotland, UK.

Erosion and dilation

The **erosion** of a binary image *f* by a structuring element *s* (denoted $f \ominus s$) produces a new binary image $g = f \ominus s$ with ones in all locations (x,y) of a structuring element's origin at which that structuring element *s* fits the input image *f*, i.e. $g(x,y) = 1$ is *s* fits *f* and 0 otherwise, repeating for all pixel coordinates (x,y).

Erosion with small (e.g. 2×2 - 5×5) square structuring elements shrinks an image by stripping away a layer of pixels from both the inner and outer boundaries of regions. The holes and gaps between different regions become larger, and small details are eliminated:

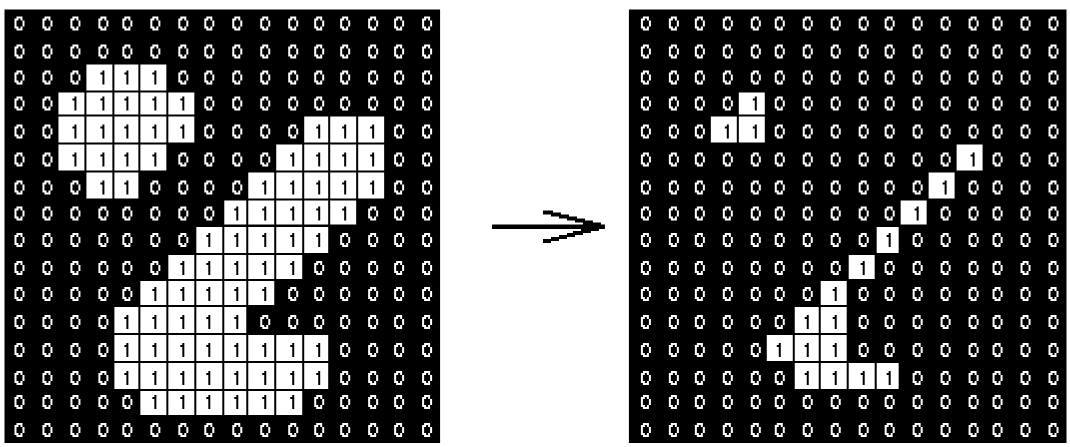


Figure 9: Erosion: a 3×3 square structuring element

Larger structuring elements have a more pronounced effect, the result of erosion with a large structuring element being similar to the result obtained by iterated erosion using a smaller structuring element of the same shape. If s_1 and s_2 are a pair of structuring elements identical in shape, with s_2 twice the size of s_1, then

$f \ominus s_2 \approx (f \ominus s_1) \ominus s_1$ [11].

Erosion removes small-scale details from a binary image but simultaneously reduces the size of regions of interest, too. By subtracting the eroded image from the original image, boundaries of each region can be found: $b = f - (f \ominus s)$ where f is an image of the regions, s is a 3×3 structuring element, and b is an image of the region boundaries.

The **dilation** of an image f by a structuring element s (denoted $f \oplus s$) produces a new binary image $g = f \oplus s$ with ones in all locations (x,y) of a structuring element's orogin at which that structuring element s hits the the input image f, i.e. $g(x,y) = 1$ if s hits f and 0 otherwise, repeating for all pixel coordinates (x,y). Dilation has the opposite effect to erosion -- it adds a layer of pixels to both the inner and outer boundaries of regions.

The holes enclosed by a single region and gaps between different regions become smaller, and small intrusions into boundaries of a region are filled in:

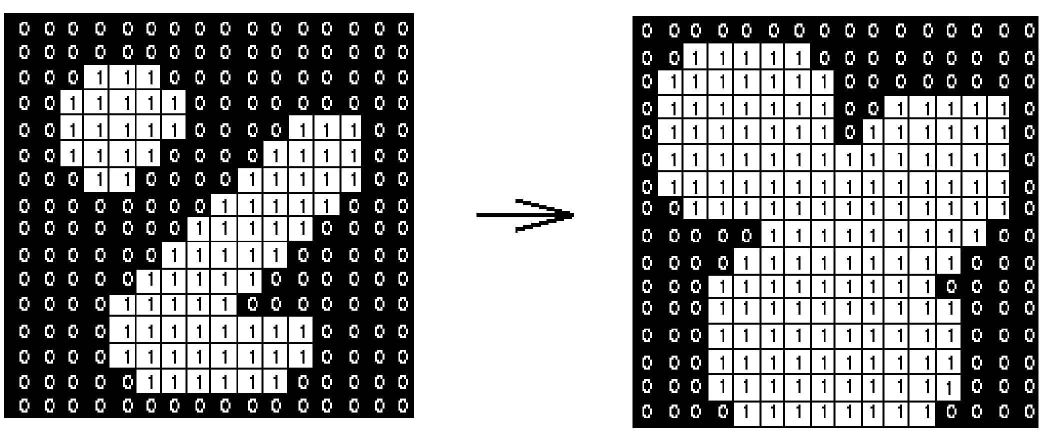


Figure 10: Dilation: a 3×3 square structuring element

Results of dilation or erosion are influenced both by the size and shape of a structuring element[10]. Dilation and erosion are *dual* operations in that they have opposite effects. Let f^c

denote the complement of an image f, i.e., the image produced by replacing 1 with 0 and vice versa. Formally, the duality is written as

$f \oplus s = f^{c} \ominus s_{rot}$

where s_{rot} is the structuring element s rotated by 180°. If a structuring element is symmetrical with respect to rotation, then s_{rot} does not differ from s. If a binary image is considered to be a collection of connected regions of pixels set to 1 on a background of pixels set to 0, then erosion is the fitting of a structuring element to these regions and dilation is the fitting of a structuring element (rotated if necessary) into the background, followed by inversion of the result.

Compound operations

Many morphological operations are represented as combinations of erosion, dilation, and simple set-theoretic operations such as the **complement** of a binary image:

$f^{c}(x,y) = 1$ if $f(x,y) = 0$, and $f^{c}(x,y) = 0$ if $f(x,y) = 1$,

the **intersection** $h = f \cap g$ of two binary images f and g:

$h(x,y) = 1$ if $f(x,y) = 1$ and $g(x,y) = 1$, and $h(x,y) = 0$ otherwise,

and the **union** $h = f \cup g$ of two binary images f and g:

$h(x,y) = 1$ if $f(x,y) = 1$ or $g(x,y) = 1$, and $h(x,y) = 0$ otherwise:

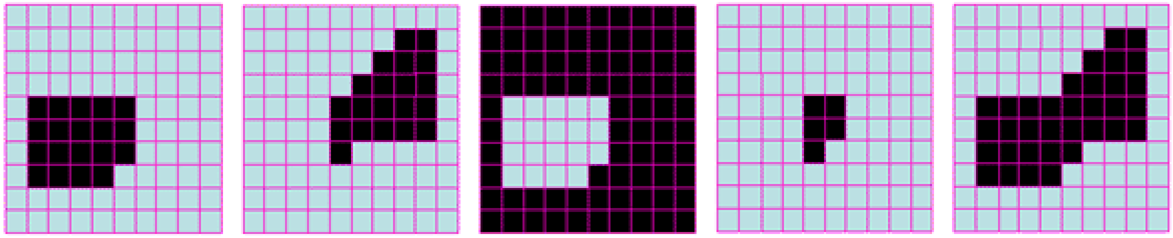
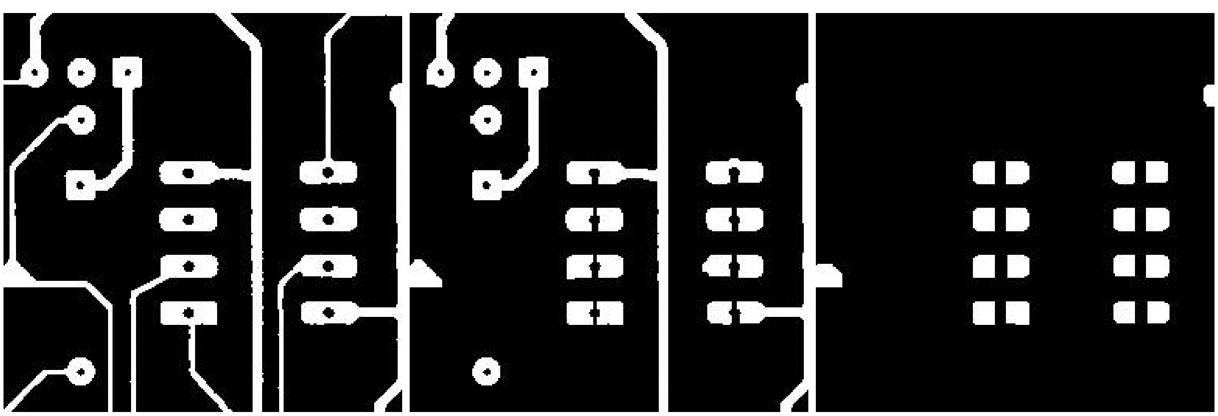


Figure 11: Set operations on binary images: from left to right: a binary image f, a binary image g, the complement f^c of f, the intersection $f \cap g$, and the union $f \cup g$.

The **opening** of an image f by a structuring element s (denoted by $f \circ s$) is an erosion followed by a dilation:

$f \circ s = (f \ominus s) \oplus s$

Opening is so called because it can open up a gap between objects connected by a thin bridge of pixels. Any regions that have survived the erosion are restored to their original size by the dilation:



Binary image f $\qquad$ $f \circ s$ (5×5 square) $\qquad$ $f \circ s$ (9×9 square)

Figure 12: Results of opening with a square structuring element

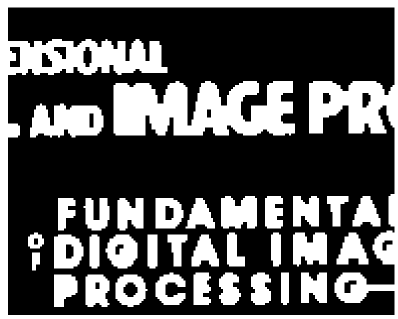
Opening is an **idempotent** operation: once an image has been opened, subsequent openings with the same structuring element have no further effect on that image:

$(f \circ s) \circ s) = f \circ s.$

The **closing** of an image f by a structuring element s (denoted by $f \bullet s$) is a dilation followed by an erosion:

$f \bullet s = (f \oplus s_{\mathrm{rot}}) \ominus s_{\mathrm{rot}}$

Binary image



Closing: a 2×2 square structuring element

Figure 13: Morphological Closing

In this case, the dilation and erosion should be performed with a rotated by 180° structuring element. Typically, the latter is symmetrical, so that the rotated and initial versions of it do not differ.

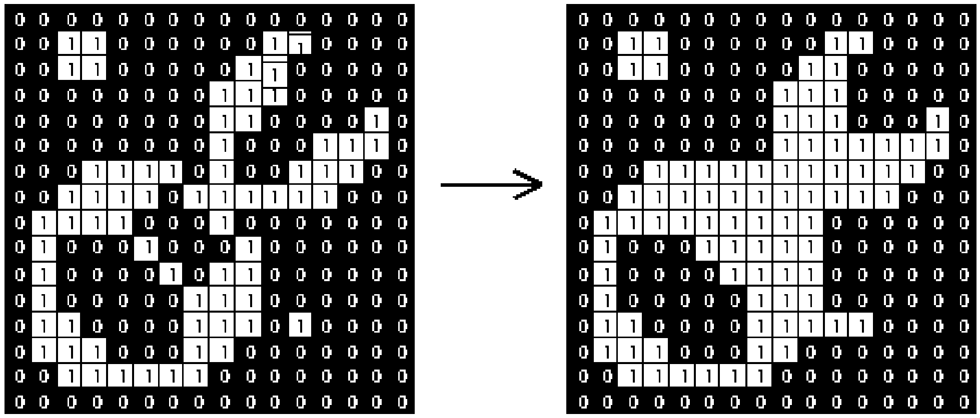


Figure 14: Closing with a 3×3 square structuring element

Closing is so called because it can fill holes in the regions while keeping the initial region sizes. Like opening, closing is idempotent: $(f \bullet s) \bullet s = f \bullet s$, and it is dual operation of opening (just as opening is the dual operation of closing):

$$f \bullet s = (f^c \circ s)^c; \quad f \circ s = (f^c \bullet s)^c.$$

In other words, closing (opening) of a binary image can be performed by taking the complement of that image, opening (closing) with the structuring element, and taking the complement of the result.

The **hit and miss transform** (see also HIPS2 web page) allows to derive information on how objects in a binary image are related to their surroundings. The operation requires a matched pair of structuring elements, $\{s_1, s_2\}$, that probe the inside and outside, respectively, of objects in the image:

$$f \circledast \{s_1, s_2\} = (f \ominus s_1) \cap (f^c \ominus s_2).$$

Binary image

Hit and miss transform:

an elongated 2×5 structuring element

Figure 15: Hit and miss trasformation

A pixel belonging to an object is preserved by the hit and miss transform if and only if s_1 translated to that pixel fits inside the object AND s_2 translated to that pixel fits outside the object. It is assumed that s_1 and s_2 do not intersect, otherwise it would be impossible for both fits to occur simultaneously.

It is easier to describe it by considering s_1 and s_2 as a single structuring element with 1s for pixels of s_1 and 0s for pixels of s_2; in this case the hit-and-miss transform assigns 1 to an output pixel only if the object (with the value of 1) and background (with the value of 0) pixels in the structuring element exactly match object (1) and background (0) pixels in the input image. Oterwise that pixel is set to the background value (0).

The hit and miss transform can be used for detecting specific shapes (spatial arrangements of object and background pixel values) if the two structuring elements present the desired shape, as well as for thinning or thickening of object linear elements.

Morphological filtering of a binary image is conducted by considering compound operations like opening and closing as filters. They may act as filters of shape. For example, opening with a disc structuring element smooths corners from the inside, and closing with a disc smooths corners from the outside. But also these operations can filter out from an image any details that are smaller in size than the structuring element, e.g. opening is filtering the binary image at a scale defined by the size of the structuring element. Only those portions of the image that fit the structuring element are passed by the filter; smaller structures are blocked and excluded from the output image. The size of the structuring element is most important to eliminate noisy details but not to damage objects of interest.

So here in our program we are going to use Morphological Closing where we will first use dilation to eliminate the gaps between pixels and lastly we will erode the frame again to make the bag size to normal size again. In this way our abandoned object being the same color as the background is still being detected without any pixel breaking between it.

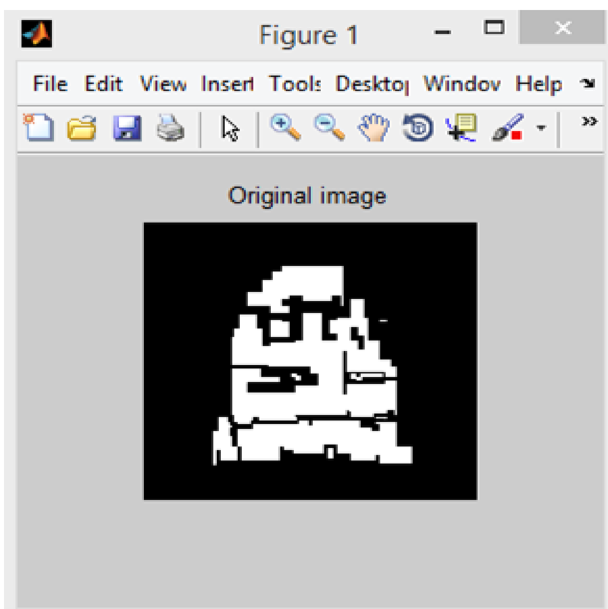


Figure 16: Image of Background Subtracted bag without Morphology

When the background color matches with the color of the abandoned object, the pixels of the matched region remains undetected. That is why we can see the image broken after background subtraction.

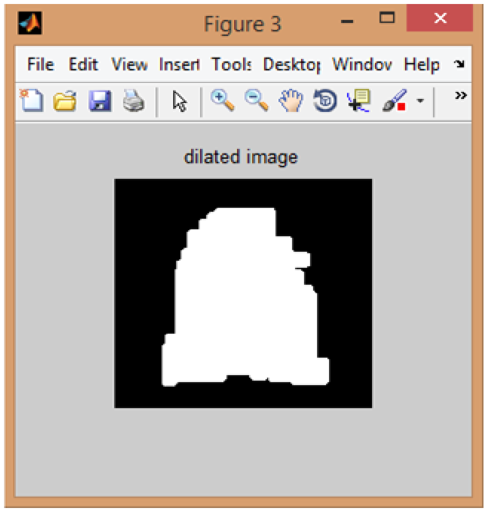


Figure 17: Dilated Image of the bag

The broken image is dilated in order to fill the gaps between the pixels.

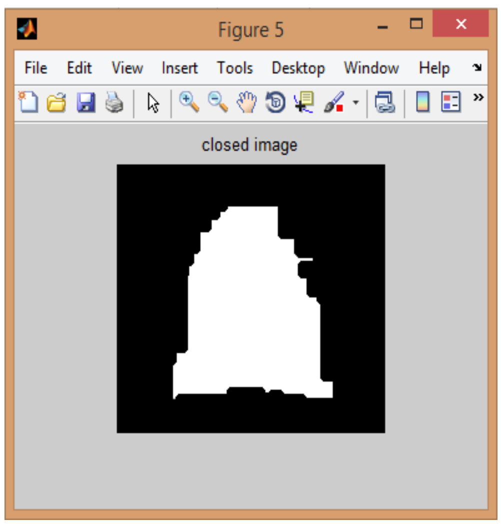


Figure 18: Morphological closing on the Bag

Since we have done dilation to fill up the gaps between the pixels, it also increases the total size of the image. That is why the dilated image is eroded to bring it back to the original size. The total process is known as morphological closing.

2.2.5 Motion-Based Human Body detection with Kalman Filter

Detection of moving objects and motion-based tracking are important components of many computer vision applications, including activity recognition, traffic monitoring, and automotive safety. The problem of motion-based object tracking can be divided into two parts:

1. detecting moving objects in each frame
2. Associating the detections corresponding to the same object over time

The detection of moving objects also uses a background subtraction algorithm based on Gaussian mixture models. Morphological operations are applied to the resulting foreground mask to eliminate noise. Finally, blob analysis detects groups of connected pixels, which are likely to correspond to moving objects.

The association of detections to the same object is based solely on motion. The motion of each track is estimated by a Kalman filter. The filter is used to predict the track's location in each frame, and determine the likelihood of each detection being assigned to each track[1] [8].

Track maintenance becomes an important aspect of this example. In any given frame, some detections may be assigned to tracks, while other detections and tracks may remain unassigned. The assigned tracks are updated using the corresponding detections. The unassigned tracks are marked invisible. An unassigned detection begins a new track.

Each track keeps count of the number of consecutive frames, where it remained unassigned. If the count exceeds a specified threshold, our system assumes that the object left the field of view and it deletes the track.

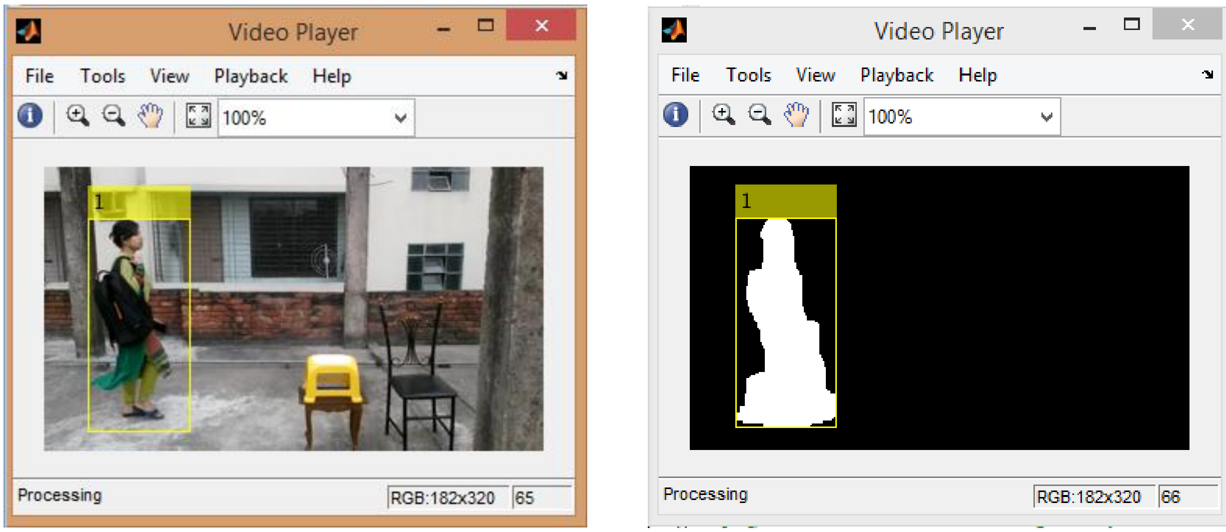


Figure 19 (a): Body Detection with Kalman Filter of person 1

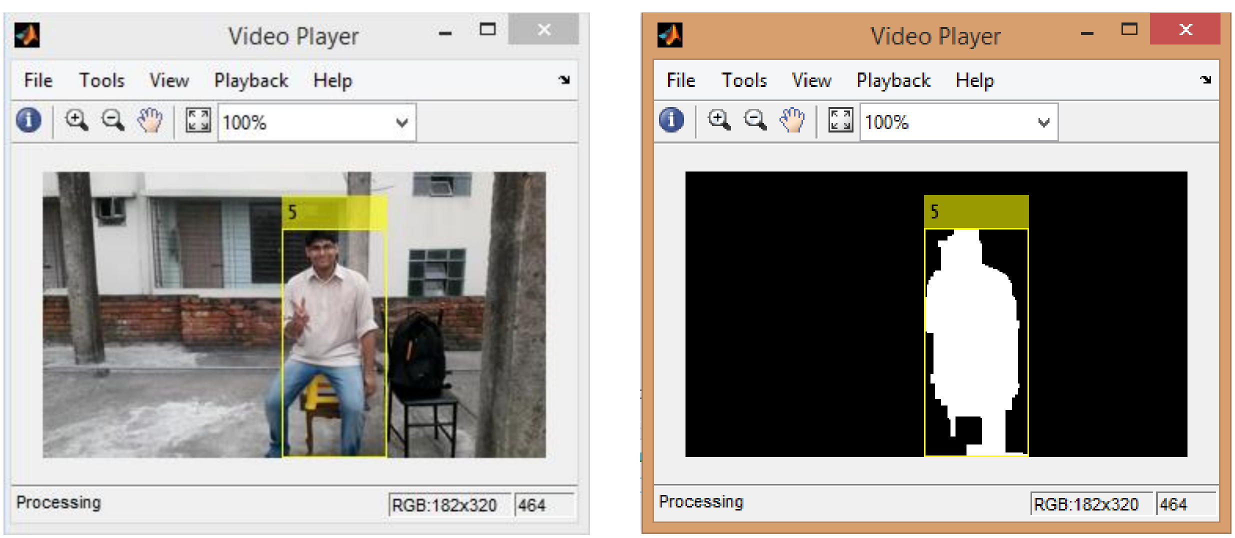
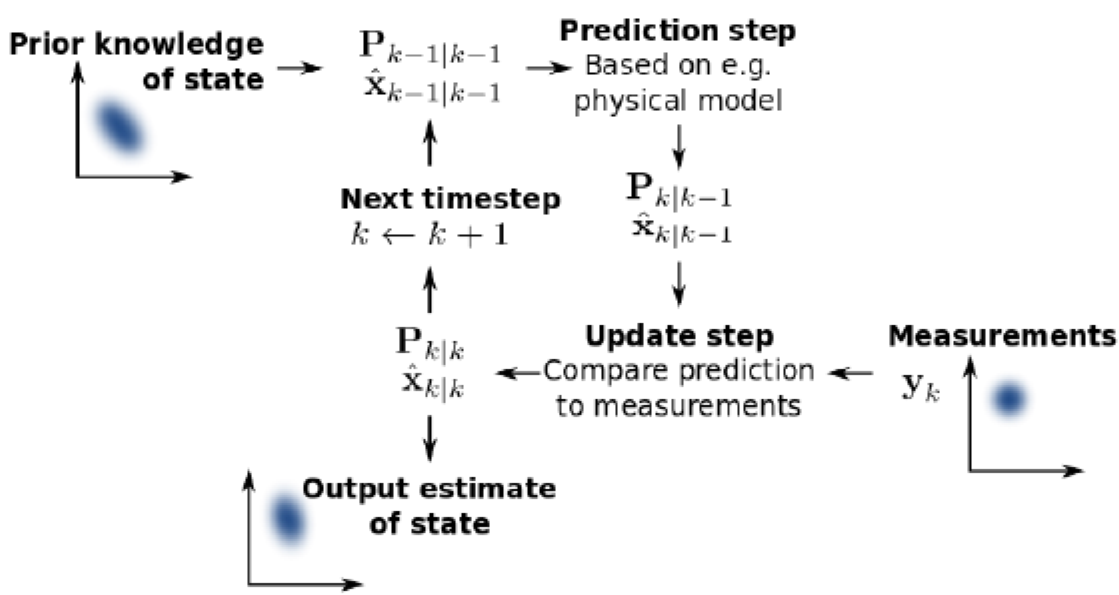


Figure 19 (b): Body Detection with Kalman Filter of person 2

The Kalman filter is designed for tracking. In our system we used it to predict a physical object's future location, to reduce noise in the detected location, or to help associate multiple physical objects with their corresponding tracks. A Kalman filter object can be configured for each physical object for multiple object tracking. To use the Kalman filter, the object must be moving at constant velocity or constant acceleration.

The Kalman filter algorithm involves two steps, prediction and correction (also known as the update step). The first step uses previous states to predict the current state. The second step uses the current measurement, such as object location, to correct the state. The Kalman filter implements a discrete time, linear State-Space System.

The Kalman filter keeps track of the estimated state of the system and the variance or uncertainty of the estimate. The estimate is updated using a state transition model and measurements. $\hat{\mathbf{x}}_{k|k-1}$ denotes the estimate of the system's state at time step k before the k-th measurement *yk* has been taken into account; $\mathbf{P}_{k|k-1}$ is the corresponding uncertainty.



One of the important lacking of Kalman filter is that the object must be moving at constant velocity or constant acceleration but the movement of the pedestrian in public places like streets, malls etc. are not constant in fact the movement is very much random. On the other hand our algorithm does not constant movement of the objects (pedestrians) rather it is functional any king of random movement, but in this case the accuracy is slightly compromised.

In our system when a body is detected it starts tracking the body and waits for a new object to appear just after the presence of the body and keeps the track of the object unalarmed waiting for the body to leave. Now, just after the leaving of the body our system waits more 45 frames for that particular body to return and checks every frame for the body. Just after the 45 frame of the absence of the body the object get red alarm even if there is/are persons of body(s).

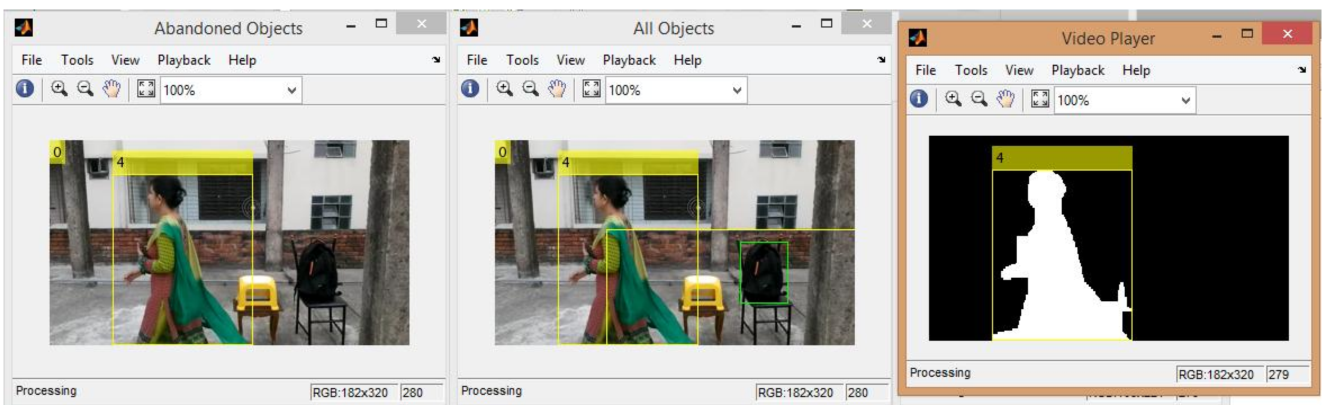


Figure 20: Sequence of detecting the carrier of the bag

In this figure the detected object did not alert with red box due to the presence of the person from whom it was detached and it is waiting for the person to leave the frame.

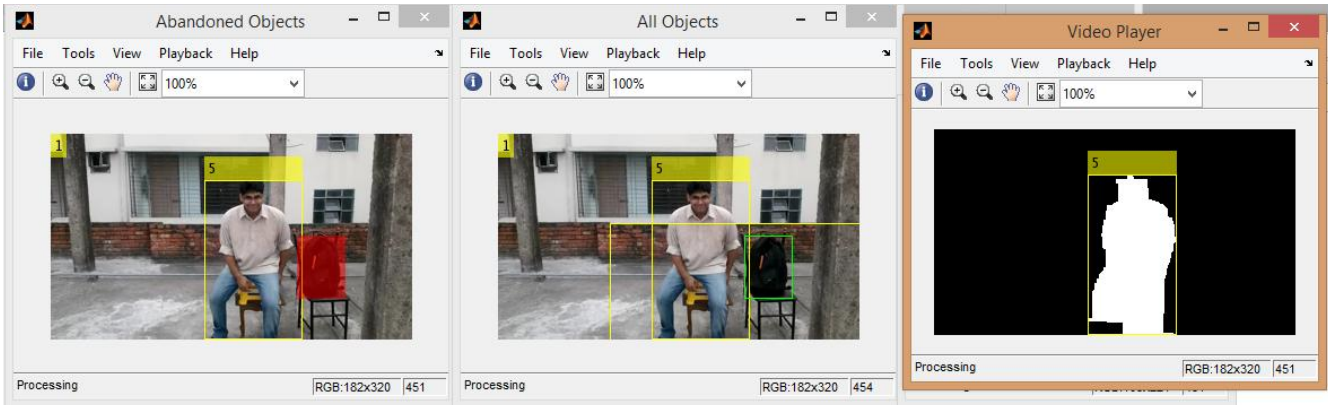


Figure 21: Detection of the bag in front of new person.

In this figure the detected object is show a red box alarm as the person in the frame is not its carrier.

3. Result and Discussion

In order to test our project, we had to take multiple videos. All the videos are of people walking with bags and leaving it on the ground, chair or table. All the videos were made by us and only one is from internet. These videos are of different frame size to check whether or system works on different frame size or not. All the videos are more than 300 frames as we need to check the abandonment of a back after each 45 frames.

3.1 Table for accuracy on Detection

We have done many experiments in our work to find the accuracy of our program. In our system detection varies on the lighting chances, movement of the camera, number of object present in the frame and the number of human body present in each frame.

3.1.1 Detection of Object before 45 frame(s)

Video No.	True positive (Correctly Identified)	True negative (Correctly rejected)	False negative (Incorrectly Rejected)	False positive (Incorrectly Identified)
1	5	3	0	1
2	9	5	1	0
3	8	4	1	1
4	7	3	0	0
5	3	6	1	2

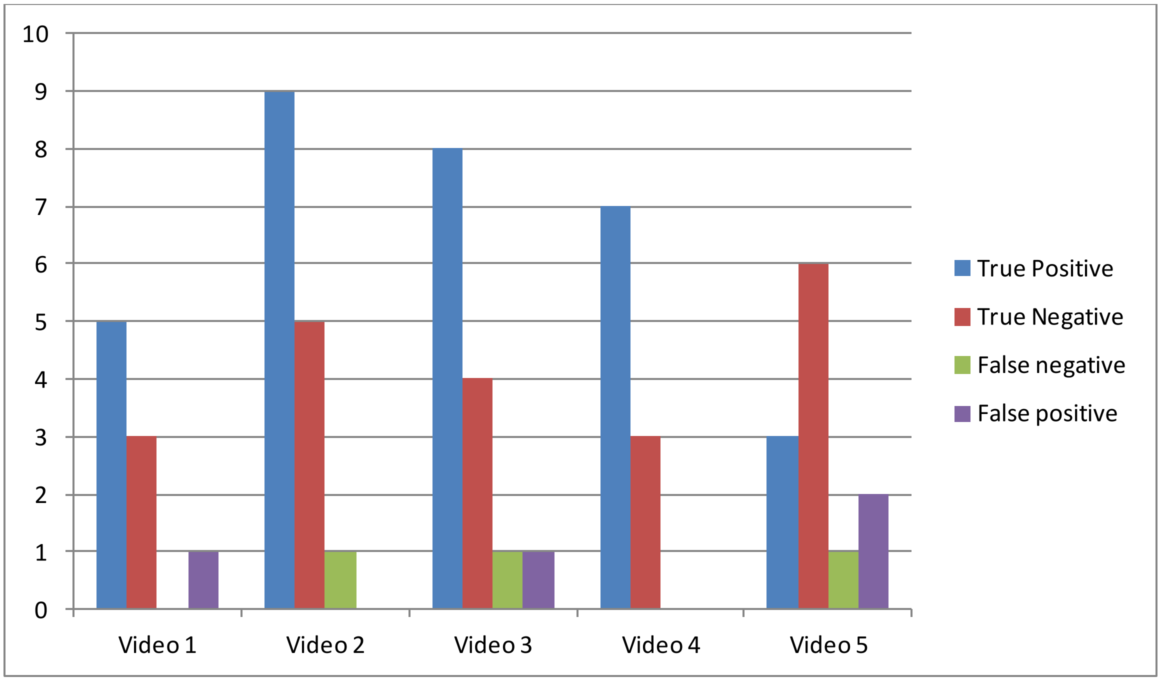


Figure 22: Chart of Accuracy of the Detection of Object before 45 frame(s)

3.1.2 Raising red box alarm for stationary abandoned object after 45 frames in ROI

Video No.	True positive (Correctly Identified)	True negative (Correctly rejected)	False negative (Incorrectly Rejected)	False positive (Incorrectly Identified)
1	1	3	0	0
2	2	2	0	0
3	3	2	0	1
4	1	0	1	3
5	1	1	2	5

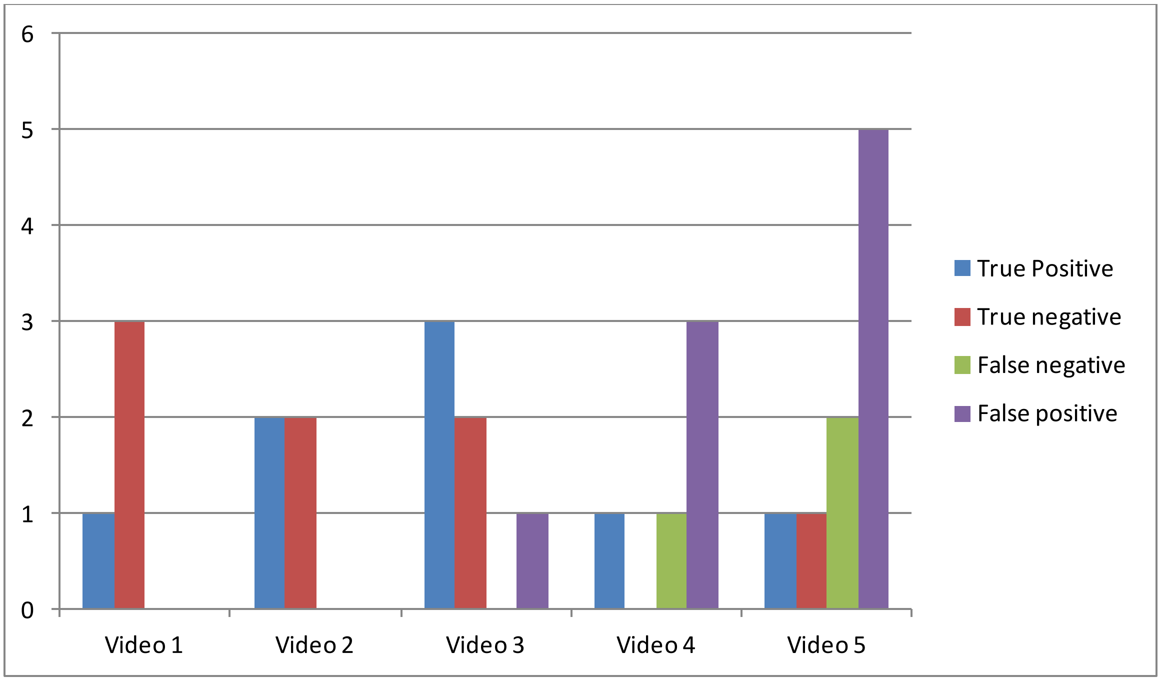


Figure 23: Chart of the Accuracy of Raising red box alarm for stationary abandoned object after 45 frames

3.1.3 Detection of Human Body

Video No.	True positive (Correctly Identified)	True negative (Correctly rejected)	False negative (Incorrectly Rejected)	False positive (Incorrectly Identified)
1	2	1	0	0
2	9	4	0	0
3	6	1	1	1
4	4	2	0	0
5	11	5	3	3

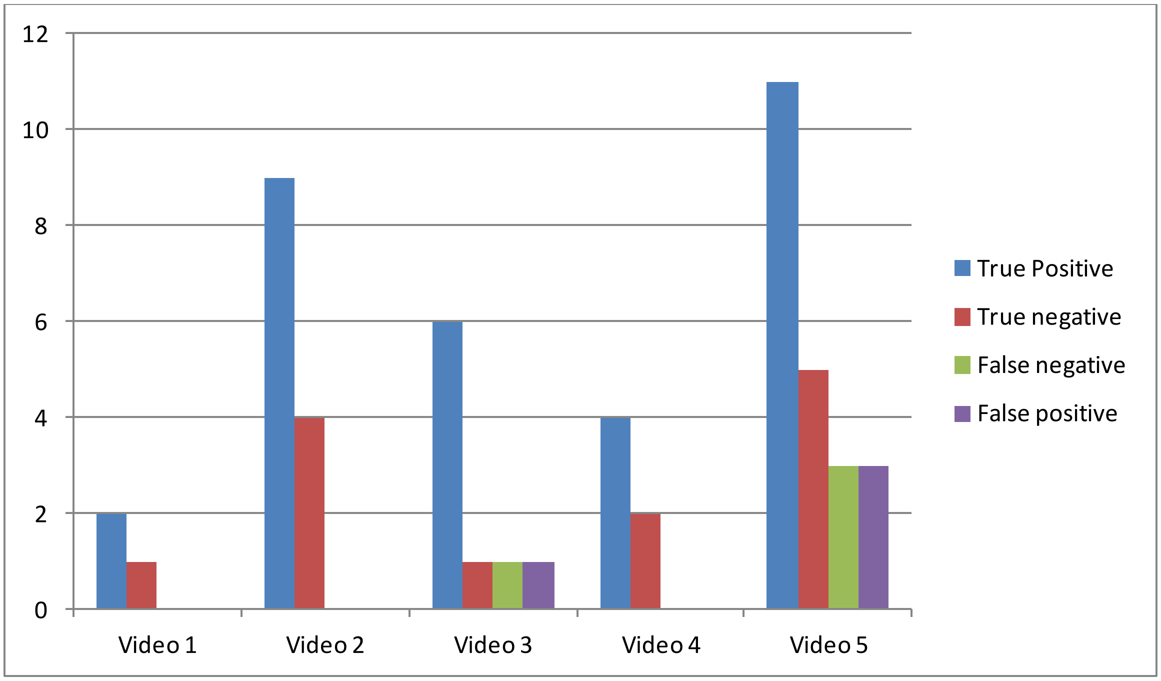


Figure 24: Chart of the accuracy of the Detection of Human Body

3.1.4 Accuracy of Detection of Abandoned Object's body (in % percentage)

Video No.	True positive (Correctly Identified)	True negative (Correctly rejected)	False negative (Incorrectly Rejected)	False positive (Incorrectly Identified)
1	87	83	13	17
2	88	81	12	19
3	83	79	17	21
4	89	82	11	18
5	71	70	29	30

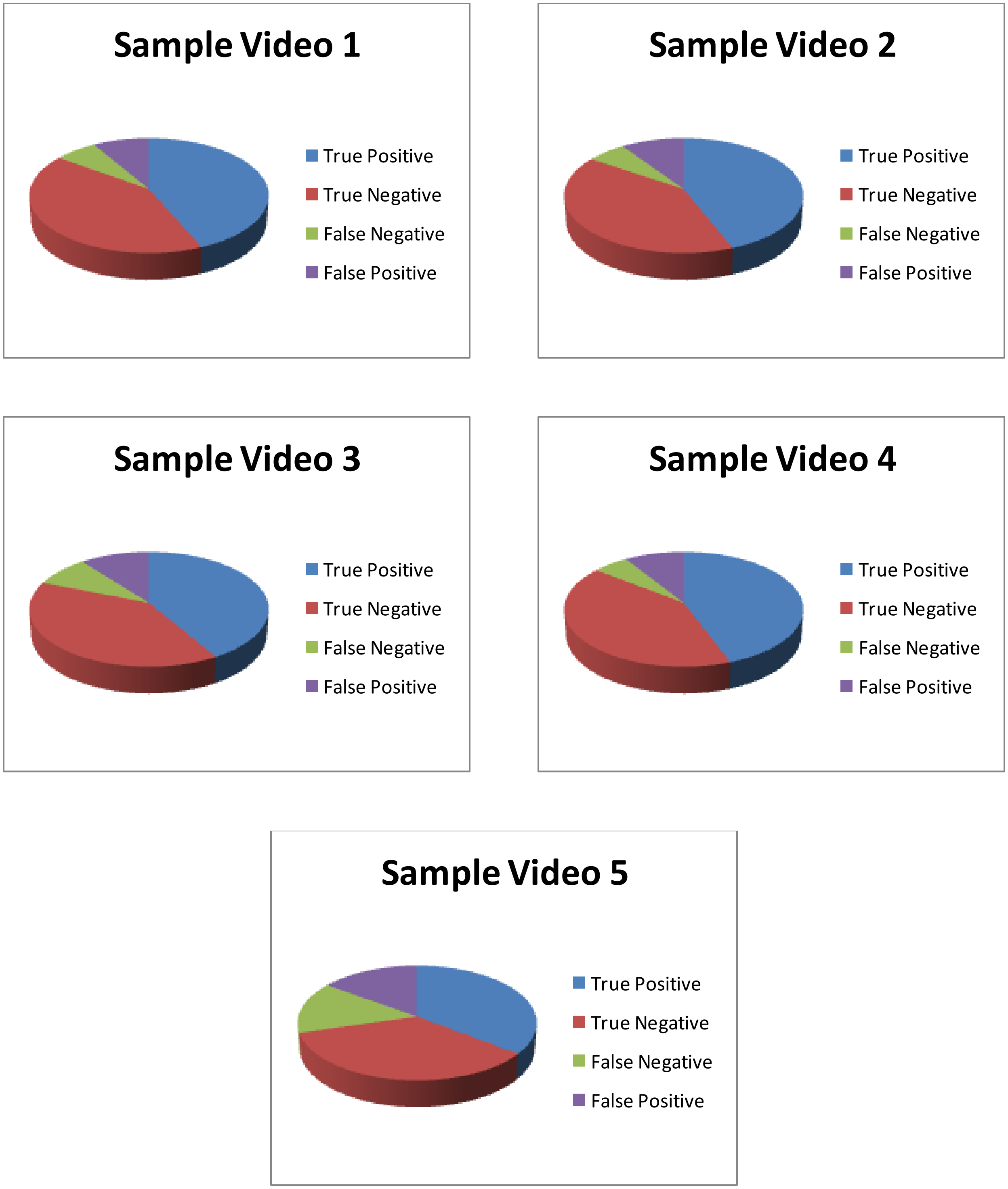


Figure 25: Chart of accuracy of Detection of Abandoned Object's body (in % percentage)

3.2 Time Analysis of detecting of Abandoned Object after Carrier leaves frame

Video No.	Total time carrier was inside the frame (minute)	Total time the bag was abandoned (minute)	Total time person was outside the frame (minute)
1	2	29	22
2	3	27	27
3	5	29	20
4	4	27	26
5	3	28	15

Table for person carrying the bag. If a person leave a bag for more than 25 minutes (45 frames) and does not return inside the frame to retrieve it then the bag will be suspected and announced as an abandoned object. So, in the following 5 cases each time video surveillance camera detected abandoned object after 25 minutes.

In our research, we worked with videos for short duration. Our result was then converted for videos that has long duration. We considered 45 frames = 5 seconds = 25 minutes.

3.3 Sequence of Output of our Video Surveillance System

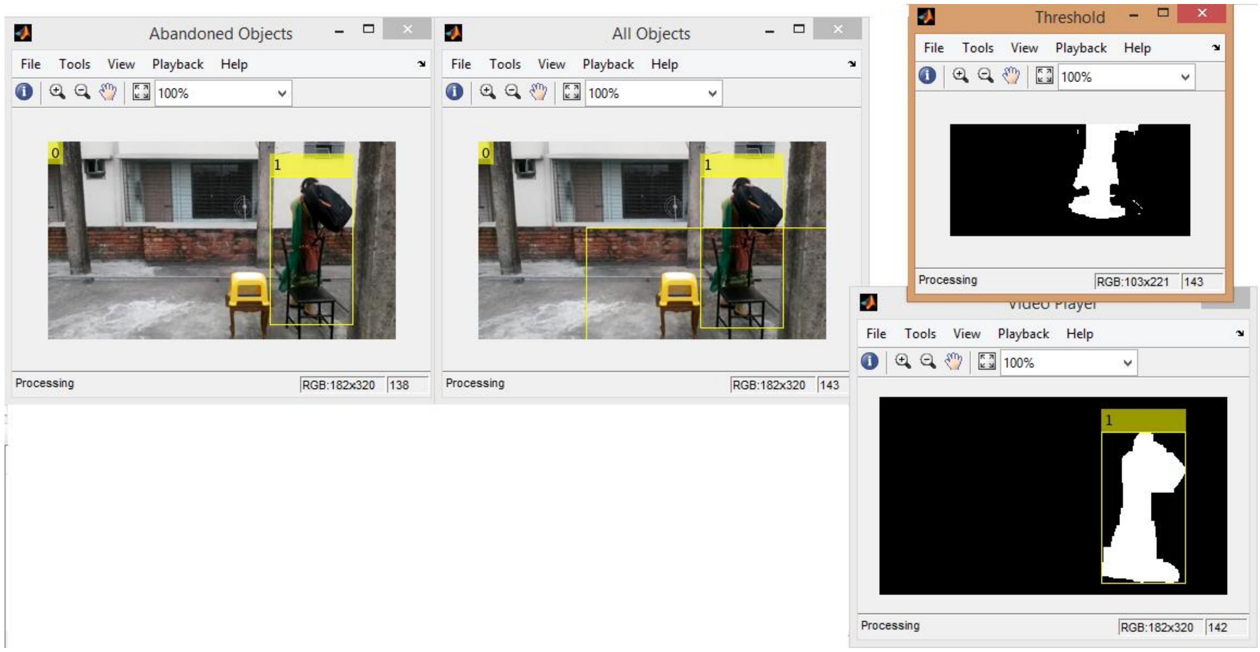


Figure 26 (a): Person abandoning the bag

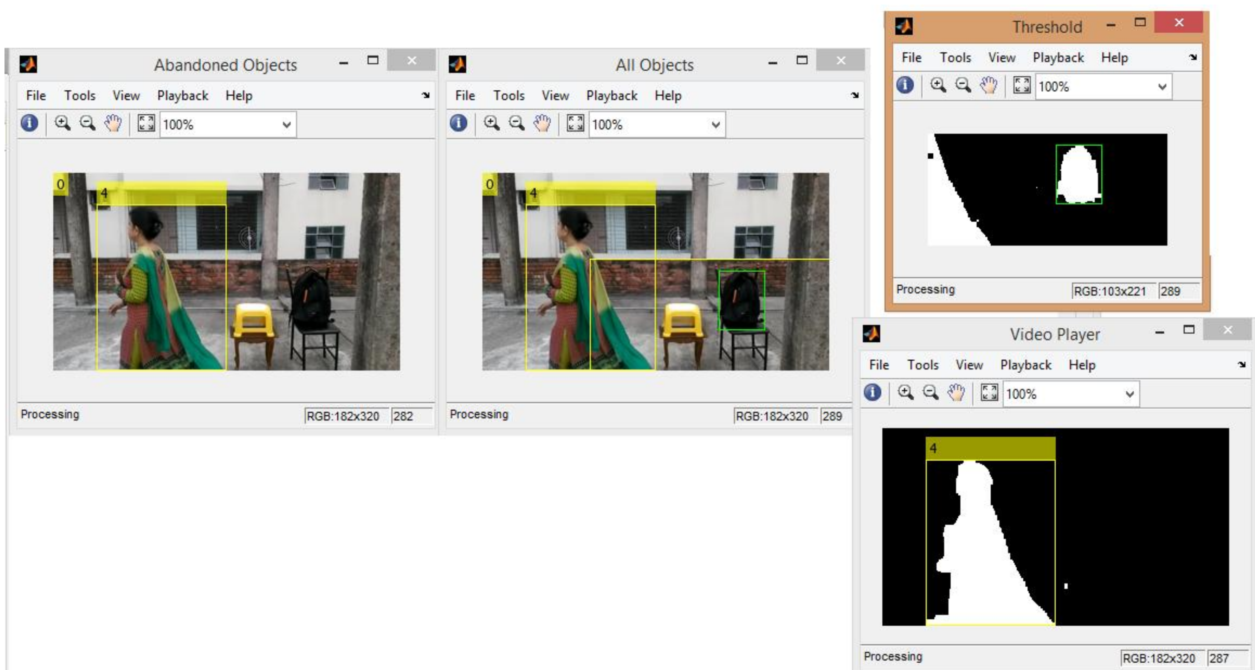


Figure 26 (b): Person leaving frame after abandoning the bag

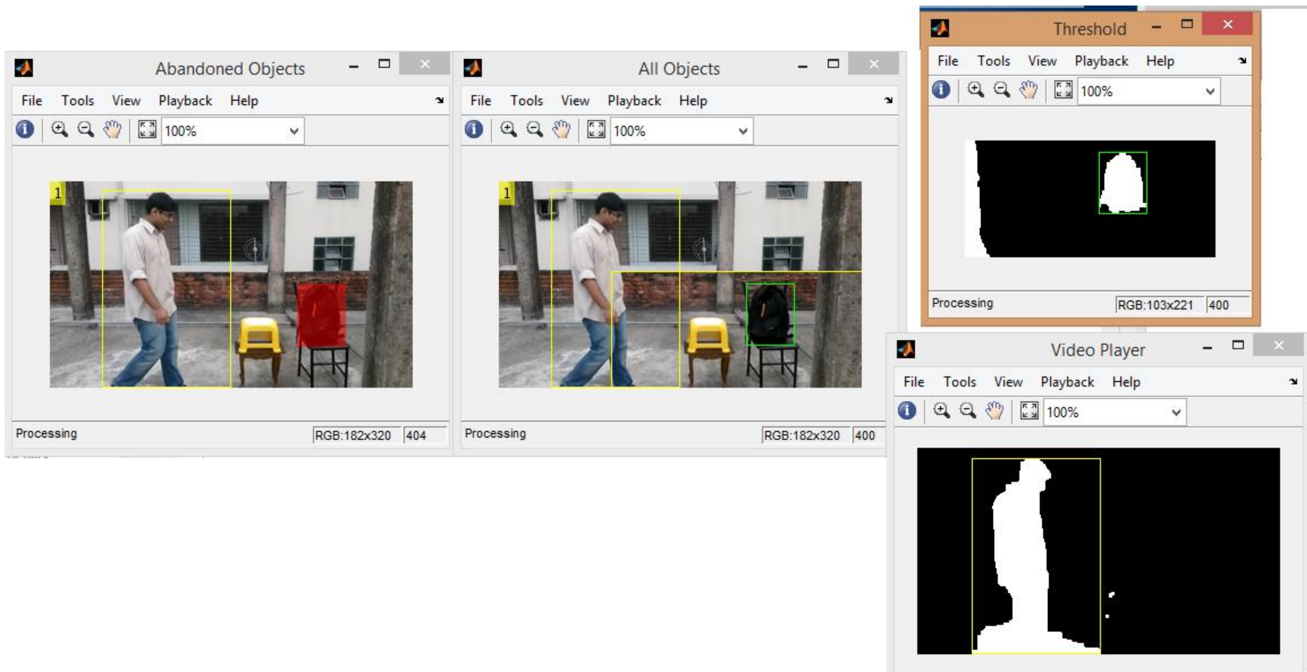


Figure 26 (c): The bag is showing abandoned by Red Box Alarm and new person entered the frame

3.4 Limitation and Challenges

While working on this system, we have faced some difficulties and problems and had to go through some limitations which are described below:

3.4.1 Moving Camera

The most significant limitation of our project is that, it does not work with moving security cameras. It is because we are segmenting and detecting object using dynamic foreground detection process and it is not possible to detect the foreground for moving cameras.

3.4.2 Overlapping objects

Sometimes when two or more objects are very close to one another then our system fails to identify all the objects separately. In that case it thinks that all the objects (very close to each other) as a single object. We can solve this problem if we decrease the threshold value then the rate of error will increase as the system will detect unwanted segments and identify them as moving objects.

3.4.3 Stationary Body

The problem with Kalman Filter is that it cannot identify a stationary body. The system work only if there is moving objects to track to.

4. Conclusion and Future Works

In this chapter we represent the conclusion of our work. Further work is also analyzed suggesting various ways through which our present research may continue.

4.1 Conclusion

Our thesis has described an efficient system for security surveillance with high accuracy of detecting multiple abandoned object. Our work also keeps track of the carrier of the object and shows alarm when the carrier leaves the frame in this way the false alarm has been minimized. In conclusion we would like to say that implementing our system would reduce more number of incidents like Boston bombing and provide everyone with more safe and crime free environment.

4.2 Future Works

We have already developed a system where the carrier is being tracked with the abandoned object and shows alarm when the carrier leaves the frame. But we still need more work to be done in our system.

a) We would like to track the carrier with his/her face.

b) Create a database and try to track and capture a terrorist by face Recognition.

c) We want to upgrade our work and implement it on a moving camera.

d) We want our system to work in very dull light environment.

e) We want to develop our system to detect suspicious activities.

f) Suicide bomber never abandons any package but stick the bomb in their body. We want to implement facial structural recognition to track the motive of any person.

Our target is not only to stop here but to implement more of these add-ons to our system in near future and make it more efficient and accurate security surveillance system.

5. References

5.1 Published Article based References

[1] Denman, S.,Sridharan, S.,Chandran, V. (n.d.). Abandoned Object Detection Using Multi-Layer Motion Detection, Image and Video Research Laboratory : Queensland University of Technology. Brisbane,AUS.

[2] Stauffer, C.,Grimson, W.E.L. (n.d.). Adaptive background mixture models for real-time tracking, The Artificial Intelligence Laboratory : Massachusetts Institute of Technology. Cambridge,MA.

[3] Wang, J.,Ooi, W. (n.d.). Detecting Static Objects In Busy Scenes, The Department of Computerf Science: Cornell University. Ithaca,NY.

[4] Pan, J., Fan, Q. ,Pankanti, S. (n.d.). ROBUST ABANDONED OBJECT DETECTION USING REGION-LEVEL ANALYSIS,IBM T.J. Watson Research Center, Hawthorne, NY.

[5] Bhargava, M., Chen, C., Ryoo, M.S., Aggarwal,J.K. (2008). Detection of object abandonment using temporal logic, Machine Vision and Applications (2009) 20:271–281.

[6] Tian, Y., Feris, R., Hampapur, A. (n.d.). Real-Time Detection of Abandoned and Removed Objects in Complex Environments, IBM T.J. Watson Research Center, Hawthorne, NY.

[7] Rout, R.K.(2013).A Survey on Object Detection and Tracking Algorithms, The Department of Computer Science and Engineering : National Institute of Technology Rourkela.

[8] Sarkar, S., Ali, R. (2014). Pedestrian Dynamics to Detect Automatic Crowd Formation, The Department of Computer Science and Engineering : BRAC University, Dhaka, BD.

5.2 Web Link based References:

[9] Morphology Fundamentals: Dilation and Erosion. (n.d.). Retrieved from http://http://www.mathworks.com/help/images/morphology-fundamentals-dilation-and-erosion.html?refresh=true

[10] Dilation. (n.d.). Retrieved from http://homepages.inf.ed.ac.uk/rbf/HIPR2/dilate.htm

[11] Erosion. (n.d.). Retrieved from http://homepages.inf.ed.ac.uk/rbf/HIPR2/erode.htm

[12] Morphology-based Operations. (n.d.). Retrieved from http://www.mif.vu.lt/atpazinimas /dip/FIP/fip-Morpholo.html

[13] Morphological Image Processing. (n.d.). Retrieved from https://www.cs.auckland .ac.nz/courses/compsci773s1c/lectures/ImageProcessing-html/topic4.htm

[14] Motion-Based Multiple Object Tracking. (n.d.). Retrieved from http://www.mathworks .com/help/vision/examples/motion-based-multiple-object-tracking.html

[15] Abandoned Object Detection. (n.d.). Retrieved from http://www.mathworks .com/help/vision/examples/abandoned-object-detection.html

[16] Background subtraction.(n.d).http://en.wikipedia.org/wiki/Background_subtraction