Occlusion handling in Object Tracking

A Dissertation submitted towards the
partial fulfilment of the requirement for the
award of degree of

Master of Technology
in
Signal Processing & Digital Design

Submitted by

NIKHIL BAJAJ
2K14/SPD/09

Under the supervision of

RAJESH ROHILLA
(Associate Professor, Department of ECE)

Department of Electronics & Communication Engineering
Delhi Technological University
(Formerly Delhi College of Engineering)
Delhi-110042

DELHI TECHNOLOGICAL UNIVERSITY

Established by Govt. Of Delhi vide Act 6 of 2009

(Formerly Delhi College of Engineering)

SHAHBAD DAULATPUR, BAWANA ROAD, DELHI-110042

CERTIFICATE

This is to certify that the dissertation title " **Occlusion handling in Object Tracking**" submitted by **Mr. Nikhil Bajaj, Roll. No. 2K14/SPD/09**, in partial fulfilment for the award of degree of Master of Technology in "**Signal Processing and Digital Design (SPDD)",** run by Department of Electronics & Communication Engineering in Delhi Technological University during the year 2014-2016 is a bonafide record of student's own work carried out by him in the academic session 2015-16 to the best of my belief and knowledge the matter embodied in dissertation has not been submitted for the award of any other degree or certificate in this or any other university or institute.

RAJESH ROHILLA

Supervisor

Associate Professor (ECE)

Delhi Technological University

DECLARATION

I hereby declare that all the information in this document has been obtained and presented in accordance with academic rules and ethical conduct. This report is my own work to the best of my belief and knowledge. I have fully cited all material by others which I have used in my work. It is being submitted for the degree of Master of Technology in Signal Processing & Digital Design at the Delhi Technological University. To the best of my belief and knowledge it has not been submitted before for any degree or examination in any other university.

NIKHIL BAJAJ

M. Tech. (SPDD)

2K14/SPD/09

Date: September, 2016

Place: Delhi Technological University, Delhi

ACKNOWLEDGEMENT

I owe my gratitude to all the people who have helped me in this dissertation work and who have made my postgraduate college experience one of the most special periods of my life. Firstly, I would like to express my deepest gratitude to my supervisor **Dr. RAJESH ROHILLA**, Associate Professor (ECE) for his invaluable support, guidance, motivation and encouragement throughout the period during which this work was carried out.

I am deeply grateful to **Dr. S.Indu,** H.O.D. (Deptt. Of E.C.E) for their support and encouragement in carrying out this project.I also wish to express my heart full thanks to all faculty at Department of Electronics & Communication Engineering of Delhi Technological University for their goodwill and support that helped me a lot in successful completion of this project.

Finally, I want to thank my parents, family and friends for always believing in my abilities and showering their invaluable love and support.

NIKHIL BAJAJ

M. Tech. (SPDD)

2K14/SPD/09

ABSTRACT

Object tracking is an essential component of an intelligent video surveillance system. Accurate and real-time object tracking will greatly improve the performance of object recognition, activity analysis and high-level event understanding. However, strength of kalman filter can be used to track object in adverse situation. Integrating a kalman filter within a standard tracking system allows the kalman filter is to use progressively updated features and aids in main training identity of the tracked object, and provides tracking system with an effective means. This thesis shows the possibilities of detecting the locations of unattended luggage's and object or a ball in occluded scenes in which the positions of these objects before occlusions are taken and then a estimate of that position is given as a primary input in kalman filter which is the main sole of this thesis. In this thesis, we tried different videos with fixed camera with a single object and multiple objects to see it is able to detect objects. This paper presents a robust tracking approach to handle challenges such as occlusion.

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Object detection and tracking

Object tracking is the boon for our modern world. With the help of it we can save our national property. We can prevent crime. It provide us healthy environment. Video observation is a dynamic examination theme in PC vision that tries to identify, perceive and track objects over an arrangement of pictures and it likewise makes an endeavor to comprehend and depict object conduct by supplanting the maturing old conventional technique for checking cameras by human administrators. Object recognition and following are vital and testing assignments in numerous PC vision applications, for example investigation, vehicle route and self-governing robot route. Object recognition includes finding objects in the edge of a video arrangement. Each tracking technique requires an article discovery instrument either in each edge or when the object first shows up in the video. Object tracking is the procedure of finding an object or various objects after some time utilizing a camera. The powerful PCs, the accessibility of high caliber and reasonable camcorders and the expanding requirement for robotized video examination has created a lot of enthusiasm for article following calculations. There are three key strides in video investigation, discovery intriguing moving items, following of such questions from every last casing to casing, and examination of item tracks to perceive their behavior. Therefore, the utilization of item following is applicable in the assignments of, movement based acknowledgment. Programmed identification, following, and checking of a variable number of articles are essential undertakings for an extensive variety of home, business, furthermore, modern applications, for example, security, reconnaissance, administration of access focuses, urban arranging, traffic control, and so forth. In any case, these applications were most certainly not as yet having essential influence in shopper hardware. The principle reason is that they need solid prerequisites to accomplish tasteful working conditions, specific and costly equipment, complex establishments and setup methodology, and supervision of qualified specialists. A few works have concentrated on creating programmed location and following calculations that minimize of supervision. They normally utilize a moving article work that assesses every speculative object configuration with the set of accessible identifications without to unequivocally register their information affiliation. Subsequently, an impressive sparing in computational expense is accomplished. Likewise, the probability capacity has been intended to represent boisterous, false and missing discoveries. The field of machine (PC) vision is worried with issues that include interfacing PCs with their encompassing surroundings. One such issue, reconnaissance, has a target to screen a given situation and report the data about the watched movement that is of significant hobby. In this admiration, video reconnaissance generally uses electro-optical sensors (camcorders) to gather data from nature. In a run of the mill reconnaissance framework, these camcorders are mounted in fixed positions or on skillet tilt gadgets and transmit video streams to a certain area, called observing room. At that point, the got video streams are checked on showcases and followed by human administrators. Be that as it may, the human administrators may face numerous issues, while they are observing these sensors. One issue is because of the actuality that the administrator must explore through the cameras, as the suspicious article moves between the constrained field of perspective of cameras and ought not to miss whatever other object while taking it. Consequently, observing turns out to be more testing, as the number of sensors in such a reconnaissance system increments. Thusly, reconnaissance frameworks must be computerized to enhance the execution and dispose of such administrator mistakes. In a perfect world, a mechanized reconnaissance framework ought to just require the destination of an application, in which constant elucidation and strength is required. At that point, the test is to give vigorous and continuous performing observation frameworks at an affordable cost. With the diminishing in expenses of equipment for detecting and figuring, and the expansion in the processor speeds, observation frameworks have turned out to be financially accessible, and they are currently connected to various different applications, for example, traffic checking, air terminal and bank security, and so on. In any case, machine vision calculations (particularly for single camera) are still extremely affected by numerous inadequacies, similar to impediments, shadows, climate conditions, and so on. As these expenses diminish just about every day, multi-camera organizes that use 3D data are turning out to be more accessible. In spite of the fact that, the utilization of various cameras prompts better treatment of these issues, contrasted with a solitary camera, shockingly, multi-camera observation is still not a definitive arrangement yet. There are some testing issues inside of the observation calculations, for example, foundation demonstrating, highlight extraction, following, impediment taking care of and occasion acknowledgment. Additionally, machine vision calculations are still not sufficiently strong to handle completely robotized frameworks and numerous exploration thinks about on such changes are as yet being finished. This work centers on building up a system to identify moving protests and produce dependable tracks from observation video. The issue is the majority of the current calculations chips away at the dim scale video. Be that as it may, in the wake of changing over the RGB v video edges to dark at the season of change, data misfortune occurs. The primary issue comes when foundation and the forefront both have roughly same dim qualities. At that point it is difficult for the calculation to find out which pixel is forefront pixel and which one foundation pixel. Now and again two different hues, for example, dull blue and dim violet, shading when changed over to dim scale, their dim qualities will come exceptionally close to each other, it can't be differentiated that which esteem originates from dim blue and which originates from dull violet. Nonetheless, if shading pictures are taken then the foundation and frontal area shading can be effectively differentiated. So without losing the shading data this modified foundation model will work straightforwardly on the shading casings.

1.2 Overview

In moving item location different foundation subtraction strategies accessible in the writing were reenacted. Foundation subtraction includes the supreme difference between the present picture and the reference upgraded foundation over a timeframe. A decent foundation subtraction ought to have the capacity to defeat the issue of shifting enlightenment condition, foundation mess, shadows, camouflage, bootstrapping and in the meantime movement division of forefront article ought to be done at the constant. It's difficult to get every one of these issues understood in one foundation subtraction method. So the thought was to mimic and assess their execution on different video information taken in complex circumstances. Object following is an exceptionally difficult assignment in the vicinity of variability brightening condition, foundation movement, and complex item shape, incomplete and full object impediments. Here in this postulation, modification is done to beat the issue of brightening variety and foundation disarray, for example, fake movement because of the takes off of the trees, water flowing, or flag waving in the wind. In some cases object following includes following of a solitary intrigued object and that is done utilizing standardized connection coefficient and overhauling the layout. On adding to a system to distinguish moving protests and produce solid tracks from observation video. In the wake of setting up an essential framework that can serve as a stage for further programmed following research, the topic of variety in separations between the camera and the items in different parts of the scene (object profundity) in observation recordings are taken. An input based answer for consequently learn the separation variety in static-camera video scenes is actualized in light of item movement in different parts of the scene. It gives more centers towards the examination of discovery and following of articles in video reconnaissance. The observation framework is the procedure of observing the conduct, exercises or other evolving data, typically individuals with the end goal of influencing, overseeing, coordinating, and ensuring. A large portion of the reconnaissance framework incorporates static camera and fixed foundation which provides some insight for the article location in recordings by foundation subtraction procedure. In reconnaissance framework three fundamental imperative steps these are item identification, object following and acknowledgment. Programmed identification, following, and checking of a variable number of articles are essential undertakings for an extensive variety of home, business, furthermore, modern applications, for example, security, reconnaissance, administration of access focuses, urban arranging, traffic control, and so forth. In any case, these applications were most certainly not as yet having essential influence in shopper hardware. The principle reason is that they need solid prerequisites to accomplish tasteful working conditions, specific and costly equipment, complex establishments and setup methodology, and supervision of qualified specialists. A few works have concentrated on creating programmed location and following calculations that minimizes the need of supervision. They normally utilize a moving article work that assesses every speculative item configuration with the set of accessible identifications without to unequivocally register their information affiliation. Subsequently, an impressive sparing in computational expense is accomplished. Likewise, the probability capacity has been intended to represent boisterous, false and missing discoveries. The field of machine (PC) vision is worried with issues that include interfacing PCs with their encompassing surroundings. One such issue, reconnaissance, has a target to screen a given situation and report the data about the watched movement that is of significant hobby. In this admiration, video reconnaissance generally uses electro-optical sensors (camcorders) to gather data from nature. In a run of the mill reconnaissance framework, these camcorders are mounted in fixed positions or on skilltilt gadgets and transmit video streams to a certain area, called observing room. At that point, the got video streams are checked on showcases and followed by human administrators

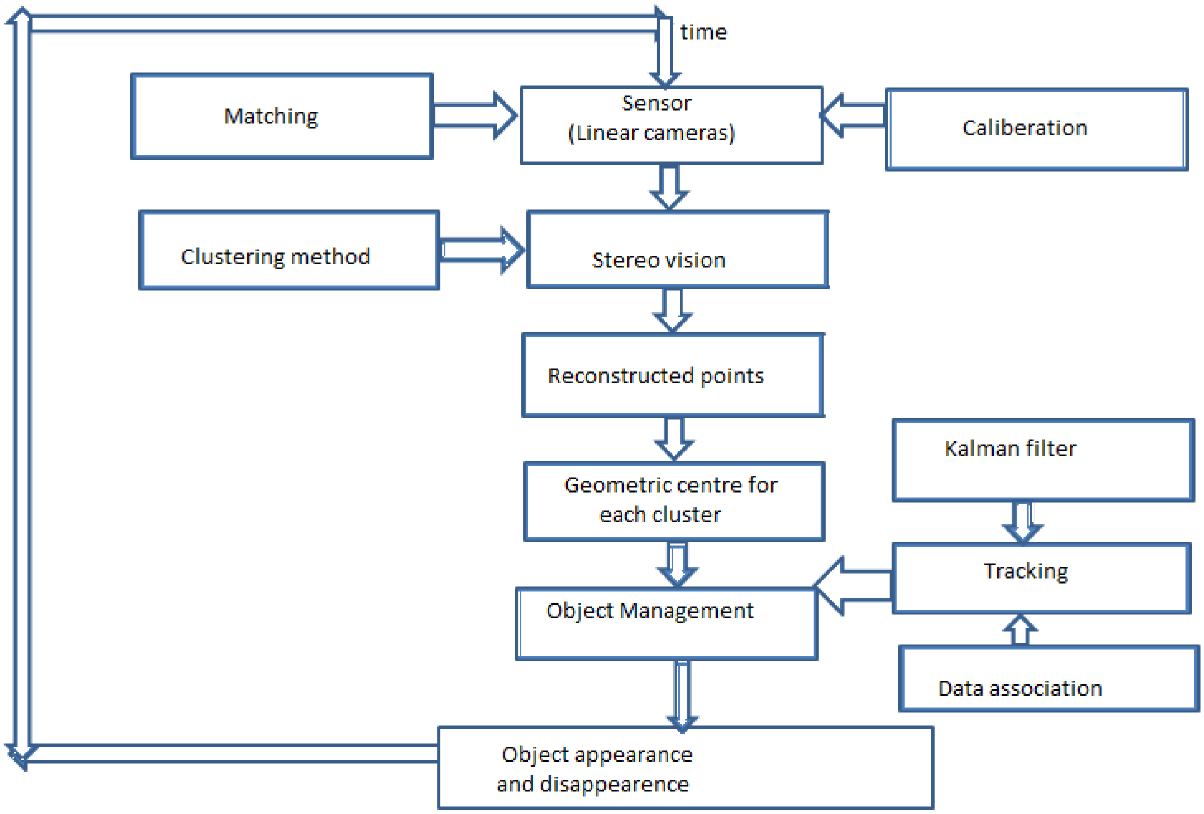


Figure 1.1: Analysis of detection and tracking approach

. In the event of video examination there are three key steps: identification of intriguing moving item, following of such protests from edge to casing and examination of articles tracks to perceive their conduct. Next it comes video division it implies detachment of items from thesensor foundation. It likewise comprises of three critical steps: object recognition, object following and question acknowledgment. In this work it is given more center towards the examination video investigation and video division segment.A typical automated single camera surveillance system usually consists of three main parts, which can be listed as moving object detection, object tracking and event recognition. In my problem it is to solve an automatic moving target detection and tracking details. The process of automatic tracking of objects begins with the identification of moving objects. An improved background subtraction method in conjunction with a novel yet simple background model to achieve very good segmentation is used. Once the moving pixels are identified, it is necessary to cluster these pixels into regions, which is referred as blobs, so that pixels belonging to a single object are grouped together. Single moving objects are often incorrectly separated into two or more sub regions because of lack of connectivity between pixels usually occurs due to occlusion from other objects

1.3 Motivation

In the wake of contemplating the writing, it is found that distinguishing the item from the video grouping furthermore track the item it is a truly difficult assignment. Object following can be a period devouring procedure because of measure of information that is contained in the video. From the writing review it is found that there are numerous foundation subtraction calculation exits which work efficiently in both indoor and open air observation framework. Julio et al. [3] has proposed a foundation demonstrating method and utilized another calculation to recognize shadowed district. Be that as it may, the shadow evacuation procedure is an overhead for item following calculation. It will be better if the shadow can be uprooted at the season of the frontal area object recognition calculation by outlining an efficient calculation, which can appropriately order the frontal area question and foundation expelling false frontal area pixel from discovery. At that point there will no additional calculation required for shadow identification and evacuation. Video observation is the most dynamic examination theme in PC vision for people what's more, vehicles. Here the point is to add to a clever visual reconnaissance framework by re-putting the age old custom technique for observing by human administrators. The inspiration in doing is to plan a video reconnaissance framework for movement discovery, what's more, question following. The territory of mechanized reconnaissance frameworks is as of now of massive enthusiasm due to its suggestions in the field of security. Reconnaissance of vehicular traffic and human exercises offers a setting for the extraction of significant data, for example, scene movement and traffic insights, object classification, human identification, abnormality identification, and also the examination of collaborations between vehicles, between people on the other hand in the middle of vehicles and people. An extensive variety of exploration conceivable outcomes is open in connection to video observation and tracking

1.4 Objective

This theory plans to enhance the execution of article recognition and following by contributing initially to two parts (a) movement division (b) object tracking. Programmed following of objects can be the establishment for some intriguing applications. A precise and efficient following ability at the heart of such a framework is fundamental for building larger amount vision-based knowledge. Following is definitely an inconsequential undertaking given the non-deterministic nature of the subjects, their movement, and the picture catch process itself. The goal of video following is to partner target objects in sequential video outlines. The affiliation can be particularly difficult when the items are moving quick with respect to the edge rate from the past segment it is found that there are numerous issues in identifying of an item and following of articles furthermore acknowledgment for fixed camera system. The objective of the work in this proposition is twofold:

1. To set up a framework for programmed division and following of moving objects in stationary camera video scenes, which might serve as an establishment for more elevated amount thinking errands and applications

2. To make significant upgrades in generally utilized calculations. At last, the point is to demonstrate to perform identification and movement based following of moving articles in a video from a stationary camera.

Hence the primary destinations are:

- To investigate division calculation to distinguish the items.
- To investigate some following strategy for following the single items and various objects.

1.5 Thesis Organization

The rest of the thesis is organized as follows:

Chapter2: The literature surveys that have been done during the research work has been discussed here. It also provides a detailed survey of the literature related to motion detection and object tracking. Discussion about the existing and some new methods for detection and tracking of objects are done. In this chapter existing methods are discussed and also examined. This chapter presents the methodology and implementation of some existing and experimental results subsequently.

Chapter 3: This chapter discusses about the background concepts related to this project work. The chapter also discusses object segmentation in image sequences, background modeling and tracking approaches. The architecture and block diagram of tracking flow systems are also explained in this chapter.

Chapter 4: This chapter provides concluding comments those can be made to the project.

Chapter 5: In this we simulated all MATLAB results through kalman filter.

Chapter 6: conclusion and future work.

CHAPTER 2

LITERATURE SURVEY

2.1 Introduction

The examination directed so far for article discovery and following items in video reconnaissance framework are discussed in this part. The arrangement of difficulties plot above traverse a few areas of examination and the lion's share of pertinent work will be investigated in the up and coming sections. In this segment, just the agent video observation frameworks are talked about for better comprehension of the key idea. Following is the procedure of object of enthusiasm inside of a grouping of edges, from its first appearance to its last. The sort of article and its portrayal inside of the framework relies on upon the application. Amid the time that it is available in the scene it might be impeded by different objects of hobby or fixed impediments inside of the scene. A following framework ought to have the capacity to foresee the position of any impeded articles.

Object following frameworks are commonly outfitted towards reconnaissance application where it is craved to screen individuals or vehicles moving around a region. There are two area ways to deal with the following issue, top-down and another is base up. Top-down strategies are objective arranged and the thefts of following frameworks are outlined in this way. These normally include some kind of division to find area of enthusiasm, from which protests and components can be extricated for the following framework. Base up react to boost and have as per watched changes. The top-down methodology is most famous strategy for creating reconnaissance framework. Framework has a typical structure comprising of a division step, an identification step

According to the portrayal in Chapter 1, object following has a considerable measure of utilization in the genuine world. Be that as it may, it has numerous innovative lacunas still exist in the strategies for foundation subtraction. In this segment, some past works is discussed for edge difference that utilization of the pixel-wise differences between two casing pictures to remove the moving areas, Gaussian blend model in light of foundation model to recognize the article what's more, finally foundation subtraction to distinguish taking so as to move areas in a picture the difference in the middle of current and reference foundation picture in a pixel-by-pixel, what's more, past works accomplished for the foundation demonstrating. After the discovery situation is over, following part is finished. Once the intriguing objects have been recognized it is helpful to have a record of their development after some time.

So following can be defined as the issue of assessing the direction of an item as the article moves around a scene. It is important to know where the item is in the picture at every moment in time. On the off chance that the items are persistent recognizable and their sizes or movement does not change after some time, and then following is not a difficult issue. By and large observation frameworks are required to watch vast region like airplane terminals, shopping centers. In these situations, it is impractical for a solitary camera to watch the complete zone of interest since sensor determination is finite and structures in the scene confine the obvious region. Accordingly reconnaissance of wide ranges requires a framework with the capacity to track objects while watching them through various cameras. Be that as it may here no discussion about various camera system is finished. Lipton et al. [5] proposed outline difference that utilization of the pixel-wise differences between two edge pictures to separate the moving districts. In another work, Stauffer and Grim son et al. [6] proposed a Gaussian blend model in view of foundation model to recognize the article. Liu et al. [7], proposed foundation subtraction to distinguish moving areas in a picture by taking the difference in the middle of current and reference foundation picture in a pixel-by-pixel. Collins et al. [8], added to a cross breed strategy that joins three-outline differencing with a versatile foundation subtraction model for their VSAM (Video Surveillance and Monitoring) venture. Desa and Salih et al [9], proposed a blend of foundation subtraction and casing difference that enhanced the past consequences of foundation subtraction and casing difference. Sugandi et al. [10], proposed another strategy for article recognition utilizing outline difference on low determination picture. Julio cezarET al. [3] has proposed a foundation model, and consolidates a novel strategy for shadow discovery in dim scale video successions. Satoh et al. [11], proposed another strategy for article following utilizing piece coordinating calculation in view of PISC picture. Sugandi et al. [12], proposed following procedure of moving persons utilizing camera fringe increase sign relationship picture. Beymer and konolige et al. [2],1999 proposed in stereo camera based article following, use kalman filter for foreseeing the articles position and speed in x-2 measurement. Rosals and sclaroff et al.,1999[19] proposed utilization of stretched out kalman filter to gauge 3D direction of an article from 2D movement.

In article location technique, numerous analysts have built up their strategies. Liu et al., 2001 proposed foundation subtraction to identify moving locales inan picture by taking the difference in the middle of current and reference foundation picture in a pixel-by-pixel. It is amazingly delicate to change in element scenes inferred from lighting and unessential occasions and so forth. In another work, Stauffer and Grimson, 1997 [20] proposed a Gaussian blend model in light of foundation model to identify the object. Lipton et al., 1998 proposed outline difference that utilization of the pixel-wise differences between two casing pictures to separate the moving districts. This technique is exceptionally versatile to element situations, however for the most part makes a lackluster display of extricating all the important pixels, e.g., there might be openings left inside moving substances. All together to overcome burden of two-edges differencing, at times three-outlines differencing is utilized. Case in point, Collins et al., 2000 [9] added to a cross breed technique that consolidates three-outline differencing with a versatile foundation subtraction model for their VSAM (Video Surveillance and Monitoring) venture. The half and half calculation effectively sections moving districts in video without the imperfections of fleeting differencing and foundation subtraction. Desa and Salih, 2004 [16] proposed a mix of foundation subtraction and casing difference that enhanced the past aftereffects of foundation subtraction and casing difference. In item following system, this article will portray more about the district based following. District based following calculations track objects agreeing to varieties of the picture districts comparing to the moving items. For these calculations, the foundation picture is kept up progressively and movement locales are typically identified by subtracting the foundation from the present picture. Wren et al., 1997 [11] investigated the utilization of little blob components to track a solitary human in an indoor environment. In their work, a human body is considered as a blend of a few blobs separately speaking to different body parts, for example, head, middle and the four appendages. The pixels having a place with the human body are allocated to the different body part's blobs. By following every little blob, the moving human is effectively followed. McKenna et al., 2000 [4] proposed a versatile foundation subtraction technique in which shading and inclination data are joined to adapt to shadows and inconsistent shading signs in movement division. Following is then performed at three levels of reflection: areas, individuals, and gatherings. Every area has a bouncing box and locales can consolidation and split. A human is made out of one or more districts assembled together under the state of geometric structure imperatives on the human body, furthermore, a human gathering comprises of one or more individuals assembled together.

Cheng and Chen, 2006 [23] proposed a shading and a spatial component of the article to recognize the track object.In the interim, the shading highlights removed is mean and standard quality of every item. Cozy et al., 2007 [32] proposed the shading dispersion of the article as perception model. The likeness of the articles estimation utilizing Bhattacharya separation. The low Bhattacharya separation compares to the high similitude. To defeat the related issue depicted over, this article proposed another strategy for article location utilizing outline difference on low determination picture Sugandi et al., 2007,[10] object following utilizing piece coordinating calculation taking into account PISC picture Satoh et al., 2001 [29] and object identification utilizing shading and spatial data of the followed object Cheng and Chen, 2006 [23].

CHAPTER 3

PROJECT IMPLEMENTATION AND ISSUES

3.1 Introduction

Object tracking is an imperative occupation inside of the field of PC vision. Object discovery includes finding objects in casings of a video grouping. Following is the procedure of finding moving articles or different items over a timeframe utilizing a camera. In fact, following is the issue of evaluating the direction or way of an item in the picture plane as it moves around a scene. The powerful PCs, the accessibility of high caliber and cheap camcorders, and the expanding requirement for computerized video examination has enhanced interest in object tracking calculations. There are three key steps in video examination:

- Detection of intriguing moving items.
- Tracking of such protests from casing to outline.
- Analysis of item tracks to perceive their conduct.

So now the inquiry emerges here that, where object following is suitable to apply? For the most part the utilization of item tracking is suitable in the task of:

- Motion-based acknowledgment
- Computerized observation
- Video indexing
- Human-PC cooperation
- Traffic observing
- Vehicle route

Tracker allocates reliable marks to the followed objects in different edges of a video. Moreover, depending upon the following space, a tracker can likewise give object-driven data, for example, introduction, region or state of an object. Following articles can be unpredictable because of:

- Loss of data created by projection of the 3D world on a 2D picture,
- Noise in pictures
- Complex article movement
- Non-inflexible or explained nature of articles
- Incomplete and full question impediments
- Complex article shapes, scene light changes
- Real-time preparing necessities.

By forcing imperatives on the movement and appearance, items can be followed. All following calculations expect that the item movement is smooth with no unexpected changes. One can compel the item movement to be of consistent speed or a steady increasing speed in light of former data. Enormous learning about the number and the span of objects, or the article appearance and shape, can likewise be utilized to improve the issue. Various methodologies for item following have been proposed. These differ from each other in light of the way they approach the accompanying inquiries:

- Which protest representation is suitable for following?
- Which picture elements ought to be utilized?
- How ought to the movement, appearance and state of the item be demonstrated?

The responses to these inquiries rely on upon the earth in which the following is performed and the end use for which the following data is being looked for. Countless strategies have been proposed which to answer these inquiries.

3.2 Object Representation

An article is just only an element of interest. Objects can be spoken to by their shapes and appearances. For instance, floats or platforms on the ocean, fish in an aquarium, vehicles on a street, planes noticeable all around, individuals strolling on a street might be imperative to track in a specific area. So there are different representations of article shape, which is regularly utilized for following and after that addresses the joint shape and appearance representations in portraying [1] as takes after.

- **Points**-The item is spoken to by a point, which is the centroid or an arrangement of focuses. The point representation is suitable when it is given more focus on the item which involved little areas in the picture. Primitive geometric shape-Geometric shape i.e. An article shape is spoken to by rectangle, circle.
- **Primitive geometric shapes**-are more suitable for speaking to basic unbending articles and non-inflexible items. Object outline and shape Contour is the limit of an article Object outline and shape Contour is the limit of an article. The area inside the form is known as the outline of the item. These representations are suitable for following complex non unbending shapes.
- **Articulated shape models**-Articulated articles are made out of body parts that are held together by joints. For instance, the human body is a verbalized article with middle, legs, hands, head, and feet associated by joints. The relationship between the parts is represented by kinematic movement models, for instance, joint point, and so forth.
- **Skeletal model**-Object skeleton can be separated by applying average hub change to the article outline. This model is regularly utilized as a shape representation for perceiving objects. Skeleton representation can be utilized to model both verbalized and inflexible items.

Object representations.in figure- 3.1 (a) Centroid, (b) various focuses, (c) rectangular patch, (d) curved patch, (e) part-based various patches, (f) object skeleton, (g) complete object form, (h) control focuses on article shape, (i) object outline.

Correspondingly there are different approaches to speak to the appearance highlight of articles. It ought to be noticed that the shape representation can be joined with appearance representations for following. Some basic appearance representations for the situation of article following are depicted in [1] as takes after.

Probability densities of article appearance-The likelihood thickness gauges the article appearance can either be parameters, for example, Gaussian and a blend of Gaussians, for example, Parzen windows and histograms. The likelihood densities of article appearance highlights (shading, composition) can be processed from the picture locales specified by the shape models (inside district ofTemplates-Templates are formed using simple geometric shapes or silhouettes. It carries both spatial and appearance information.
Problems identified in article appearances: Templates, however, only encode the object appearance generated from a single view. Thus, they are only suitable for tracking objects whose poses do not vary considerably during the course of tracking.

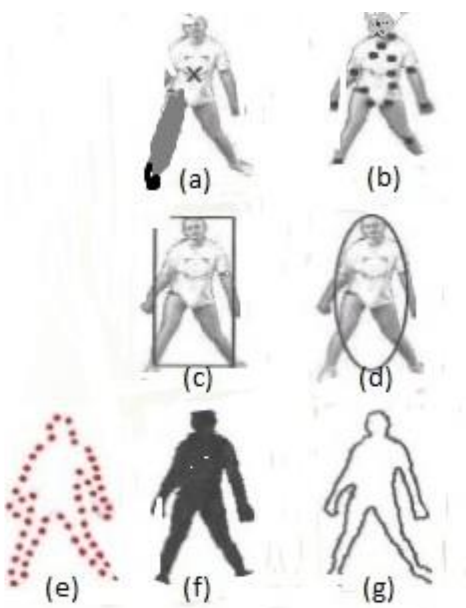


Figure 3.1: Object Representation

- Points:- figure above shows the points to be used in object representation.
- Primitive geometric shape: - shapes like rectangle, eclipse can be used to represent.
- Object silhouette and contour:-contour part defines the boundary of object. The region inside the contour is called the silhouette of contour

Active appearance models-these are generated by simultaneously modeling the object shape and appearance. Object shape is defined by a set of land marks. Each landmark, an appearance vector is stored in the form of color, texture or gradient magnitude. These models required a training phase where both the shapes & its associated appearance is learned from a set of samples.

Problems identified for Active Appearance models-Legitimately identifying articles can be an especially difficult errand, particularly since articles can have rather convoluted structures and may change fit as a fiddle, size, area and introduction over resulting video outlines.

Multi view appearance models- These models encode different views of an object. One approach to represent the different object views is to generate a subspace from the given view. Example of subspace approaches is Principal Component Analysis (PCA), Independent component Analysis (ICA).
Problems identified for Multi view appearance model: One limitation of multi-view appearance models is that the appearances in all view shave required a lot of time.

In general, there is a strong relationship between object representation and tracking algorithms. (Object representations are chosen according to the applicationdomain an oval or a shape) [1].

3.3 Algorithm used for implementation:

In algorithm representation we represent each and every step used for further representing, detecting and tracking objects. Firstly we load video and creates its Gaussian mixture model then developing its foreground and background model after that increase object detection or blob detection rate of probability by using interests points in a blob and defining its Kalman filter then carried for prediction and correction of result of blob analysis.

3.3.1. Load video: load any video for locating any object for detection and tracking. Capturing images through any video with high quality and good size is so easy because of rapid improvement in quality of capturing device with less costly but superior technology. Videos are a collection of sequential images with a constant time interval. So video can provide more information of scenarios generated by any object in any constant interval of time. Therefore, manually handling videos are quite impossible. There is need an automated devise to process these videos. In this thesis one such attempt has been made to track objects in videos.

.

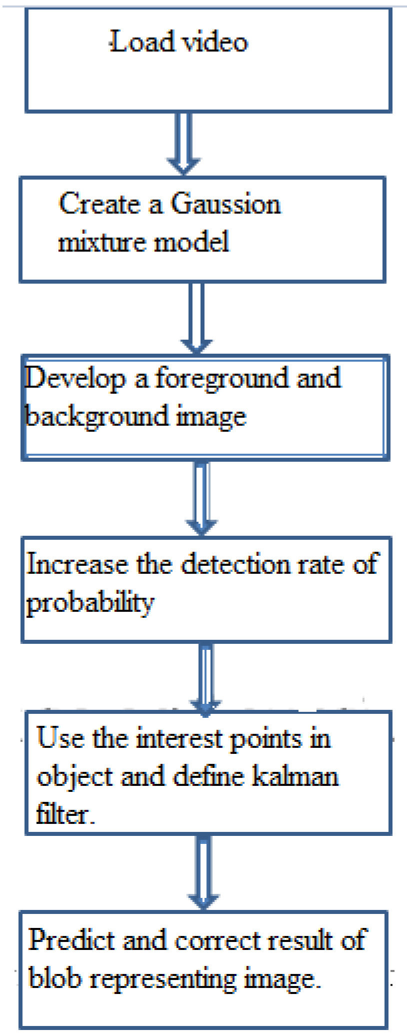


Figure 3.2 flowchart of implementation

3.3.2. Creating Gaussian mixture Image:

A Gaussian mixture model is a probabilistic model that accepts every one of the information focuses are produced from a mixture of a limited number of Gaussian distribution with unknown parameters. One can consider mixture models as summing up k-implies grouping to join data about the covariance structure of the information and also the focuses of the inert Gaussians. Mixture models are a sort of thickness model which involve various part works, generally Gaussian. These part capacities are joined to give a multimodal thickness. They can be utilized to demonstrate the shades of an item keeping in mind the end goal to perform shops, for example, ongoing shading based following and division. These shops might be made more hearty by creating a mixture model relating to foundation hues notwithstanding a forefront model, and utilizing Bayes' hypothesis to perform pixel order. Mixture models are additionally agreeable to compelling strategies for on-line adjustment of models to adapt to gradually fluctuating lighting conditions.

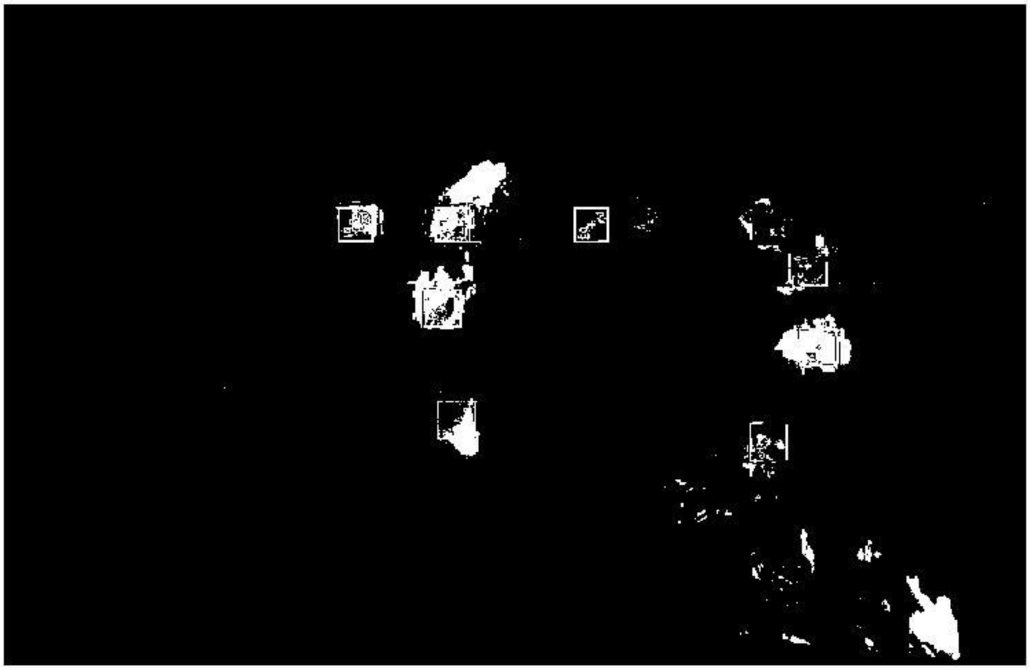


Figure 3.3 Gaussian image

3.3.3.Developing a foreground and background model:

3.3.3.1 Point detectors

Point detectors are used to find interesting points in images which have an expressive texture in their respective localities. A desirable quality of an interest point is its invariance to changes in illumination and camera viewpoint. In literature, commonly used interest point detectors include Moravecs detector, Harris detector, KLT detector, SIFT detector. Any object following calculation will contain blunders which will in the long run cause a float from the object of hobby. The better calculations ought to have the capacity to minimize this float such that the tracker is precise over the time period of the application.

3.3.3.2 Segmentation

Subsequent to preprocessing the following step is division. Division implies, separate the items from the foundation. The point of picture division calculations is to parcel the picture into perceptually comparative locales. Each division calculation addresses two issues, the criteria for a decent parcel and the strategy for accomplishing efficient apportioning. In the writing overview it has been examined different division procedures that are applicable to question following. They are, Mean movement grouping, and picture division utilizing Graph-cuts (N2.4.4 Motion Segmentation.

The first work in any observation application is to recognize the objective articles in the video outline. Most pixels in the casing have a place with the foundation and static districts, and suitable calculations are expected to distinguish singular focuses in the scene. Since movement is the key marker of target nearness in observation recordings, movement based division plans are generally utilized. An effective and straightforward strategy for approximating the foundation that empowers the location of individual moving objects in video edges is being used. These are the different division strategy which will be talked about in subtle elements. There are various propositions for the arrangement of moving article identification issue in the observation framework. In spite of the fact that there are various propositions for the arrangement of moving item discovery issue in reconnaissance frameworks, some of these strategies are discovered to be all the more encouraging by the analysts in the field. Strategies are similar to

- Frame differencing strategy to identify objects.
- Mixture of Gaussian in light of moving article identification technique.
- Background subtraction technique to recognize forefront objects.

3.3.3.3 Foreground Segmentation

Closer view division is the procedure of partitioning a scene into two classes; one is foregrounding another is foundation. The foundation is the area, for example, streets, structures and furniture. While the foundation is fixed, its appearance can be required to change after some time because of elements, for example, changing climate or lighting conditions. The closer view any component of the scene that is moving or anticipated that would move and some closer view components might really be stationary for long stretches of time, for example, stopped autos, which might be stationary for quite a long time at once. It is too conceivable that a few components of foundation might really move, for example, trees moving in a breeze. The principle ways to deal with finding frontal area objects inside in the reconnaissance framework is Foundation demonstrating or subtraction-approaching pixels contrast with a foundation model to figure out whether they are forefront or foundation. The model that can be utilized by any following calculation to speak to the intrigued article is known as target representation. That model incorporates the data of intrigued object about the shape, size and appearance in a picture. The model relies on upon the intrigued question and following calculation that are utilized. There are distinctive approaches to show an intrigued object:-

- It may define priori of interested object
- It may snapshot of interested object
- It may be decided by training sample

There are two ways of target representation

- Shape representation, e.g., centroid, rectangle, ellipse, rigid model, contours or point distribution model

- Appearance representation, e.g., template, histogram

3.3.3.4Background Subtraction

Object detection can be achieved by building are presentation of the scene called the background model and then finding deviations from the model for each incoming frame .Any significant change in an image region from the background model signifies a moving object. The pixels constituting the regions undergoing change are marked for further processing. This process is referred to as the background subtraction. There are various methods of background subtraction as discussed in the survey are Frame differencing Region-based (or) spatial information, Hidden Markov models (HMM) and Eigen space decomposition.

3.3.3.5 Image Segmentation

The aim of image segmentation algorithms is to partition the image into perceptually similar regions. Every segmentation algorithm addresses two problems, the criteria for a good partition and the method for achieving efficient partitioning. In the literature survey it has been discussed various segmentation techniques that are relevant to object tracking [1] They are, mean shift clustering, and image segmentation using Graph-Cuts (Normalized cuts) and Active contours. Object detection can be performed by learning different object views automatically from a set of examples by means of supervised learning mechanism.

3.3.3.6 Ratio histogram

A histogram is a graphical representation of the distribution of numerical data. It is an estimate of the probability distribution of a continuous variable (quantitative variable). To construct a histogram, the first step is to "bin" the range of values—that is, divide the entire range of values into a series of intervals—and then count how many values fall into each interval. The bins are usually specified as consecutive, non-overlapping intervals of a variable. The bins (intervals) must be adjacent, and are usually equal size.

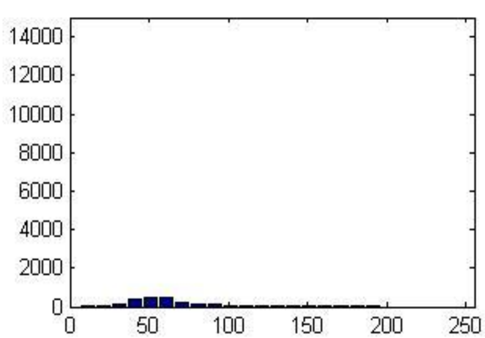


Figure 3.4: Rose image and its histogram

3.3.4 Object Detection

The execution of an observation framework impressively relies on upon its first step that is identification of the frontal area objects which don't have a place with the foundation scene. These closer view districts have a significant part in resulting activities, such as following and occasion discovery.

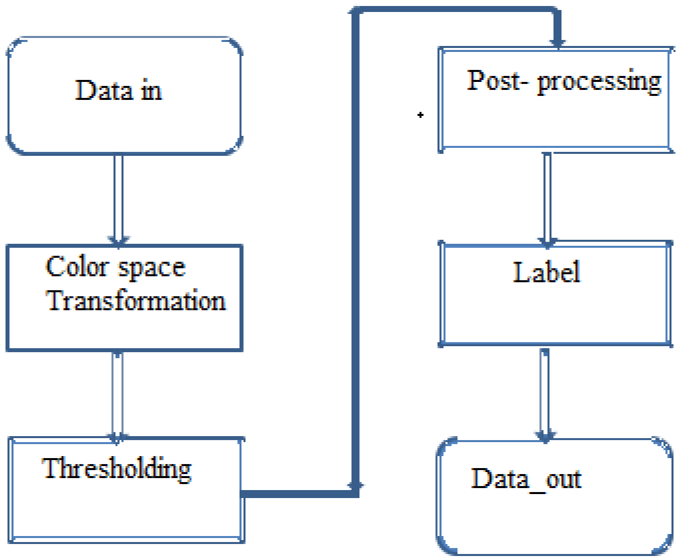


Figure 3.5: Flow of procedure

Mainly object detection method consists of two main steps. The first step is a preprocessing step including gray scaling, smoothing, and reducing image resolution and so on. The second step is filtering to remove the image noise contained in the object. The filtering is performed by applying the morphology filter such as dilation and erosion. And finally connected component labeling is performed on the filtered image

Pre-processing: In the preprocessing phase, the first step of the moving object detection process is capturing the image information using a video camera. In order to reduce the processing time, a gray scale image is used on entire process instead of the color image. The gray scale image only has one color channel that consists of 8 bits while RGB image has three color channels. Image smoothing is performed to reduce image noise from input image in order to achieve high accuracy for detecting the moving objects. The smoothing process is performed by using a median filter with m×m pixels. Here, un-stationary background such as branches and leaf of a tree as part of the background are considered. The un-stationary background often considers as a fake motion other than the motion of the object interest and can cause the failure of detection of the object. To handle this problem, the resolution of the image is reduced to be a low resolution image.

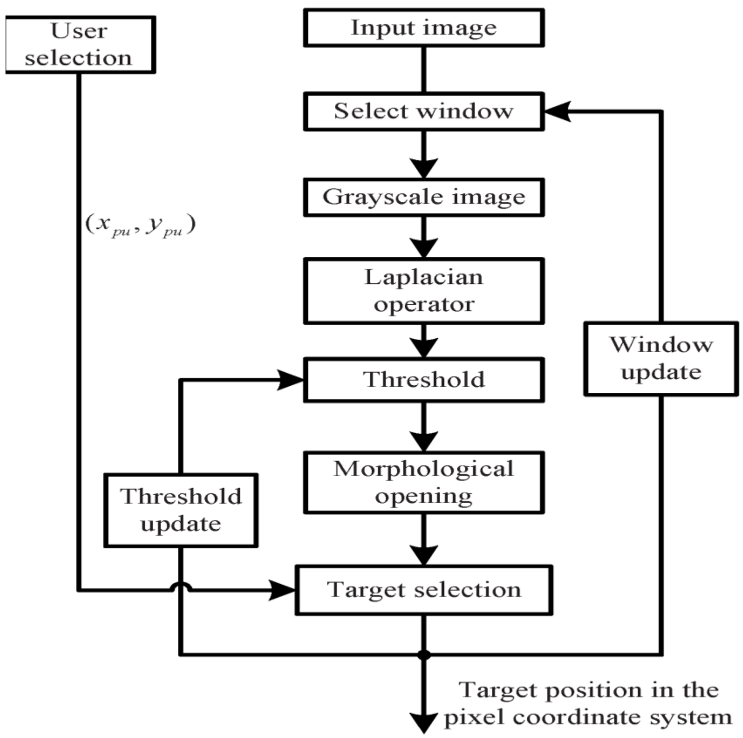


Figure 3.6: flow of object detection

3.3.4.1 Increased detection rate of probability

The goal of video detection is to partner target objects in back to back video outlines. The affiliation can be particularly difficult when the items are moving quickly in respect to the edge rate. Another circumstance that builds the many-sided quality of the issue is when followed object changes introduction after some time. For these circumstances video following frameworks more often than not utilize a movement model which portrays how the picture of the objective may change for the different conceivable movement of the items. Various methodologies for item following have been proposed. These essentially differ from each other in light of the way they approach the accompanying inquiries: Which object representation is suitable for following? Which picture elements ought to be utilized? By what method ought to the movement, appearance and state of the article be demonstrated?

The responses to these inquiries rely on upon the connection or environment in which the following is performed and the end use for which the following data is being looked for. Moving object division is basically taking into account a correlation between the data outline and a specific foundation mode, and different locales between the info and the model are marked as frontal area in view of this examination. This appraisal can be the basic casing differencing, if the foundation is static (has no moving parts what's more, is simpler to show). In any case, more intricate correlation techniques are required to fragment closer view locales when foundation scenes have dynamic parts, for example, moving tree limbs and shrubberies. In the writing, there are different calculations, which can adapt to these circumstances that will be examined in the accompanying areas.

Almost, every tracking framework begins with movement location. Movement location goes for isolating the relating moving article area from the foundation picture. The first process in the movement identification is catching the picture data utilizing a camcorder. The movement identification stage incorporates some picture preprocessing step, for example, dark scaling and smoothing, diminishing picture determination utilizing low determination picture method, outline difference, morphological operation and marking. The preprocessing steps are connected to decrease the picture clamor with a specific end goal to accomplish a higher precision of the following. The using so as to smooth system is performed middle filter. The lower determination picture is performed in three progressive casings to uproot the little or fake movement out of sight. At that point the casing difference is performed on those casings to identify the moving item developing on the scene. The next procedure is applying a morphological operation, for example, enlargement and disintegration as filtering to diminish the clamor that remaining parts in the moving article. Associated part marking is then performed to name each moving article in different name. The second stage is tracking the moving article. In this stage, a piece coordinating procedure to track just the enthusiasm moving article among the moving items developing out of sight is performed. The pieces are defined by isolating the picture outline into non-covering square parts. The pieces are made in view of PISC picture that considers the shine change in every one of the pixels of the pieces in respect to the considered pixel. The last stage is item Normalized cuts) and Active contours. Performed on those casings to identify the moving objects developing on the scene. The next procedure is applying a morphological operation, for example, enlargement and disintegration as filtering to diminish the clamor that remaining parts in the moving article. Associated part marking is then performed to name each moving article in different name. The second stage is following the moving article. In this stage, a piece coordinating procedure to track just the enthusiasm moving article among the moving items developing out of sight is performed. The pieces are defined by isolating the picture outline into non-covering square parts. The pieces are made in view of PISC picture that considers the shine change in every one of the pixels of the pieces in respect to the considered pixel.

The last stage is object identification. For this reason spatial and shading data of the followed object as the picture highlight is utilized. At that point, an element line is made to spare the components of the moving articles. At the point when the new questions show up on the scene, they will be followed and marked, and the elements of the item are removed and recorded into the line. Once a moving item is distinguished, the framework will remove the components of the article and recognize it from the identified objects in the line. A couple subtle elements of every stage are portrayed as takes after.

3.3.4.2Moving Object Detection

Execution of a robotized visual observation framework extensively relies on upon its capacity to identify moving objects in the watched environment. A consequent activity, for example, following, breaking down the movement or recognizing objects, requires a precise extraction of the foreground objects, making moving object detection a crucial part of the system. In order to decide on whether some regions in a frame are foreground or not there should be a model for the background intensities. Any change, which is caused by a new object, should be detected by this model, whereas un-stationary background regions, such as branches and leaves of a tree or a flag waving in the wind, should be identified as a part of the background. So to handle these problems a method was proposed. Location of article and following of that protest is an essential errand in the zone of PC vision application. In article discovery we find or distinguish intrigued object in back to back edges of a video document. Tracking is a procedure to find moving intrigued object or different articles in a video document or camera regarding time. Actually, in object tracking we gauge or characterize the direction or way of an intrigued object in the casing plane as it moving around the picture plane. Since of innovation expanding in computational force, accessibility of good quality and ease camcorder and the need of mechanized video framework individuals are sowing the more enthusiasm for object following calculation.

3.3.5 Object Tracking

Every tracking method requires an object detection mechanism either in every frame or when the object first appears in the video. A common approach for object detection is to use information in a single frame. However, some object detection methods make use of the temporal information computed from a sequence of frames to reduce the number of false detections. For object detection, there are several common object detection methods described in Object tracking on a very basic level involves assessing the area of a specific district in progressive edges in a video arrangement. Legitimately identifying articles can be a especially difficult errand, particularly since articles can have rather convoluted structures and may change fit as a fiddle, size, area and introduction over resulting video outlines. Different calculations and plans have been presented in the few points. The point of an article tracker is to create the direction of an object after some time by finding its position in each edge of the video. In any case, following has two definition one is in truly it is finding a moving article or different object over a timeframe utilizing a camera. Another in actually following is the issue of evaluating the direction or way of an item in the picture plane as it moves around a scene. The undertakings of recognizing the object and setting up a correspondence between the article occurrences crosswise over edges can either be performed independently or together. In the first case, conceivable article district in each casing is acquired by method for an item recognition calculation, and afterward the tracker compare objects crosswise over edges. In the last case, the article locale and correspondence is mutually evaluated by iteratively overhauling object area and district data acquired from past edges. There are different strategies for Tracking.

- **Point tracking-** Tracking can be formulated as the correspondence of detecting objects represented by points across frames. Point tracking can be divided into two broad categories, i.e. Deterministic approach and Statistical approach. Objects detected in consecutive frames are represented by points, and the association of the points is based on the previous object state which can include object position and motion. Active appearance models: Interested item can be likewise spoken to by
- **Dynamic appearance models**. These sorts of models are delivered by all the while exhibiting the article appearance and shape. When all is said in done, the article shape is described by an arranged of turning points. Like as form based representation, the notable point can live on the limit of article or, then again, they can stay inside the item shape. For each noteworthy point, an appearance vector is put away as angle greatness, surface or shading. Appearance models oblige a readiness stage where both its related appearance and the shape are picked up from a circumstance of illustration used.

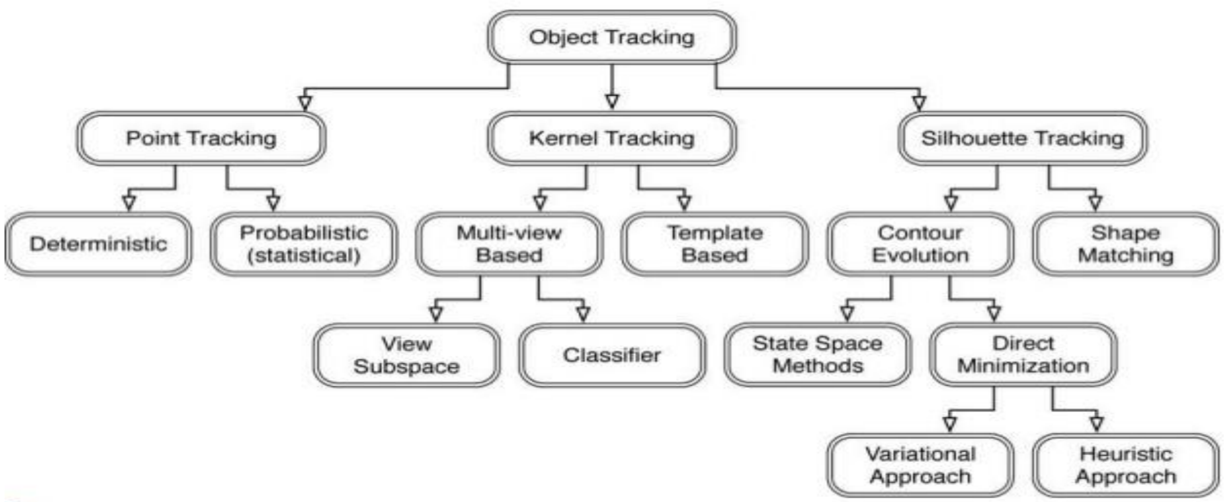


Figure 3.7: Different Tracking categories

- **Kernel tracking-** Performed by computing the motion of the object, represented by a primitive object region, from one frame to the next. Object motion is in the form of parametric motion or the dense flow field computed in subsequent frames. Kernel tracking methods are divided into two subcategories based on the appearance representation used i.e. Template and Density-based Appearance Model and Multi-view appearance model.
- **Silhouette Tracking-** It Provides an accurate shape description of the target objects. The goal of silhouette tracker is to find the object region in each frame by means of an object

model generated using the previous frames. Silhouette trackers can be divided into two categories i.e. Shape matching and Contour tracking.

Object tracking consists in estimating of the trajectory of moving objects in the sequence of images. The most important is the automation of object tracking is a challenging task. Dynamics of multiple parameters, changes representing features and motion of the objects and temporary partial or full occlusion of the tracked objects have to be considered.

3.3.5.1 Feature Selection for Tracking

It assumes an essential part to choose an appropriate element in following. So highlight choice is firmly identified with the article representation. For instance, shading is utilized as an element for histogram based appearance representations, while for form based representation, object edges is normally utilized as elements. For the most part, numerous following calculations utilize a blend of these elements. The points of interest of regular visual elements are as per the following :

- **Color:-**Color of an object is influenced by two factors. They are Spectral power distribution of the illuminant and Surface reflectance properties of the object. Different color models are RGB, L*u*v and L*a*b used to represent color.
- **Edges:-**Edge detection is used to identify strong changes in image intensities generated by object boundary. Edges are less sensitive to illumination changes compared to color features. Most popular edge detection approach is Canny Edge detector.
- **Optical Flow:-**It is a dense field of displacement vector which defines the translation of each pixel in a region. It is computed using the brightness constraint, which assumes brightness constancy of corresponding pixels in consecutive frames. Optical Flow is commonly used as a feature in motion based segmentation and tracking application.
- **Texture:-**Texture is a measure of the intensity variation of a surface which quantifies properties such as smoothness and regularity. It requires a processing step to generate the descriptors. There are various texture descriptors: Gray-Level Co-occurrence Matrices, loss texture measures, wavelets, and steerable pyramids.

For the most part components are picked physically by the client relying upon the application. The issue of programmed highlight choice has gotten significant consideration in the example acknowledgment group. Programmed highlight choice techniques can be separated into, Filter Methods and Wrapper Methods. Channel strategies attempt to choose the elements taking into account general criteria, though Wrapper techniques chooses the elements in light of the helpfulness of the elements in a specific issue space. Anything taking after computation will contain botches which will over the long haul cause a buoy from the object of pastime. The better counts should have the ability to minimize this buoy such that the tracker is exact over the time of the application.

3.3.5.2 Single Camera Object Tracking:

Till now various concepts of object tracking are being discussed. It is necessary to know how the tracking occurs in front of a single fix camera. It is very important to track properly in a single camera so that it will be easy for us to track it in multiple cameras. Whenever tracking is being done using single camera there are various challenges need to be taken care of.

- A single person may enter in to the FOV of camera.
- More than one person may enter in to the FOV of camera.
- Object carrying some goods.

First Motion Segmentation to extract moving blobs in the current frame is performed. Some blobs that are very small and are likely to be noise are deleted. The object tracking module tracks these moving objects over successive frames to generate object tracks.n article following the imperative test that needs to consider while the working a video tracker are the point at which the foundation is show up which is like intrigued object or another article which are available in the scene. This phenomenon is known as clutter.

A low resolution image is done by reducing spatial resolution of the image with keeping the image size. The low resolution image can be used for reducing the scattering noise and the small fake motion in the background because of the un-stationary background such as leaf of a tree. These noises that have small motion region will be disappeared in low resolution image. Next it comes filtering phase. In order to fuses narrow breaks and long thin gulfs, eliminates small holes, and fills gaps in the contour, a morphological operation is applied to the image. As a result, small gaps between the isolated segments are erased and the regions are merged. To extract the bounding boxes of detecting objects, connected component analysis was used. Morphological operation is performed to fill small gaps inside the moving object and to reduce the noise remained in the moving objects. The morphological operators implemented are dilation followed by erosion. In dilation, each background can be expressed as:

$$F(x,y) = \begin{cases} 1, if\ there\ is\ one\ or\ more\ pixels\ of\ the\ 8\ neighbours\ are\ 1 \\ 0, \hfill otherwise \end{cases} \quad (3.1)$$

In erosion, each object pixel that is touching a background pixel is changed into a background pixel. Erosion removes isolated foreground pixels. Erosion can be expressed as:

$$F(x,y) = \begin{cases} 0\ , if\ there\ is\ one\ more\ pixels\ of\ the\ 8\ neighbours\ are\ 0 \\ 1, \hfill otherwise \end{cases} \quad (3.2)$$

3.3.5.3 Tracking Methods:

The following calculations displayed for the most part depend on the appraisal of the intrigued object position in the video outline plane. In a particular application, it is an operational condition which is satisfactory while a building up a tracker where it can depend on instatement by client. In the constant fluctuating use of various intrigued objects, the tracker needs to utilize computerized instatement and mechanized end capacity.A direction (state) is the way that a moving item finishes space as a component of time. A direction can be portrayed scientifically either by the geometry of the way or as the position of the article after some time. It stores the genuine way of the object of interest, i.e. data about the objective in back to back outlines. We will get the all data around an objective question that in which course it moves and what is the velocity of target.

- **Point is tracking-** Tracking can be formulated as the correspondence of detecting objects represented by points across frames. Point tracking can be divided into two broad categories, i.e. Deterministic approach and Statistical approach.
- **Kernel tracking-** Performed by computing the motion of the object, represented by a primitive object region, from one frame to the next. Object motion is in the form of parametric motion or the dense flow field computed in subsequent frames. Kernel tracking methods are divided into two subcategories based on the appearance representation used i.e. Template and Density-based Appearance Model and Multi-view appearance model.
- **Silhouette Tracking-**It Provides an accurate shape description of the target objects. The goal of silhouette tracker is to find the object region in each frame by means of an object model generated using the previous frames. Silhouette trackers can be divided into two categories i.e. Shape matching and Contour tracking.

3.3.5.4 Prediction Methods:

An important part of a tracking system is the ability to predict where an object will be next frame.After the object detection is achieved, the problem of establishing a correspondence between object masks in consecutive frames should arise. Obtaining the correct track information is crucial for subsequent actions, such as object identification and activity recognition. A block matching technique is used for this purpose. Object tracking is a video processing application with a wide number of applications. Applications may include tracking particular people in a video for security reasons. for following planetary items from satellite information for cosmic studies. An object of hobby is characterized on the premise of specific application which is available within reach. An object of hobby may rely on upon the sort of utilization. For instance, in activity reconnaissance application intrigued item might be human or auto, though for satellite application intrigued article might be a planet or for gaming application it might be face of specific individual. In an object following calculation, an object of hobby is characterized on the premise of application which are available yet it can be utilized for further investigation. From past illustration, it is clear that we need to take such question as objects of hobby which object following. For instance, creature in the zoo, rocket in war, individuals in the shopping center, expanding on satellite, and so on are the case of set of intrigued article which might be the most critical to distinguish or track in a specific application. An intrigued item might be demonstrated by their appearance and shape. In article representation intrigued item can be demonstrated distinctive way that can help to video tracker. This is needed to aid the matching the tracks to detect objects and to predict the position during occlusion. There are four common approaches to predict the objects' positions:

1. Block matching

2. Kalman filters

3. Motion models

4. Particle filter`

3.3.5.5 Block Matching Method

The entire process of tracking the moving object is illustrated in the following Fig 3.5. The block matching method is well described in [4], which is applied here.

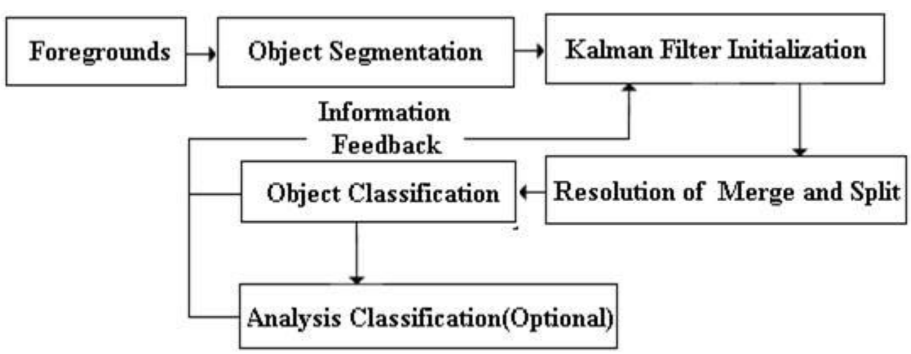


Figure 3.8: flow of tracking objects

Block matching is a technique for tracking the interest moving object among the moving objects emerging on the scene. In this article, the blocks are defined by dividing the image frame into non-overlapping square parts. The blocks are made based on peripheral increment sign correlation (PISC) image that considers the brightness change in all the pixels of the blocks relative to the considered pixel. In the Fig it shows the block in the PISC image with block size is 55 pixels. Therefore, one block consists of 25 pixels. The blocks of the PISC image in the previous frame are defined as shown in Eq. (2.3). Similarly, the blocks of the PISC image in the current frame are defined in Eq. (2.4). To determine the matching criteria of the blocks in two successive frames, the evaluation is done using a correlation value that expresses in Eq. (2.5). This equation calculates the correlation value between block in the previous frame and the current one for all pixels in the block. The high correlation value shows that the blocks are matching each other. The interest moving object is determined when the number of matching blocks in the previous and current frame are higher than the certain threshold value. The threshold value is obtained experimentally.

$$b_{np} = \begin{cases} 1, & if\ f_{np \geq f(i,j)} \\ 0, & otherwise \end{cases} \qquad (3.3)$$

$$b'_{np} = \begin{cases} 1, & if\ f_{np \geq f(i,j)} \\ 0, & otherwise \end{cases} \qquad (3.4)$$

$$corr^{n} = \sum_{P=0}^{N} b_{np} * b'_{np} + \sum_{P=0}^{N} (1 - b_{np}) * (1 - b'_{np}) \qquad (3.5)$$

The tracking method used in this article can be described as follows. The matching process is illustrated in Fig.2.6. Firstly, blocks and the tracking area are made only in the area of moving objects to reduce the processing time. The block size (block A) is made with 9x9 pixels in the previous frame. It is assumed that the object coming firstly will be tracked as the interest moving object. The block A will search the matching block in each block of the current frame by using correlation value as expressed in Eq. (3.5). In the current frame, the interest moving object is tracked when the object has maximum number of matching blocks. When that matching criteria are not satisfied, the matching process is repeated by enlarging the tracking area (the rectangle with dash line). The blacks still are made in the area of moving objects. When the interest moving object still cannot be tracked, then the moving object is categorized as not interest moving object or another object and the tracking process is begun again from the begin.

3.3.5.6 Tracked Feature Extraction:

The feature of objects extracted in the spatial domain is the position of the tracked object. The spatial information combined with the features in time domain represents the trajectory of the tracked object, so the movement and the speed of the moving objects that are tracked can be estimated. Therefore, the features of spatial domain are` very important to object identification. The bounding box defined in Eq. (2.4) is used as spatial information of moving objects. After getting the interest moving object, that is extracted by using a bounding box. The bounding box can be determined by computing the maximum and minimum value of x and y coordinates of the interest moving object according to the following equation:

$$B^{i}_{min} = \{(x^{i}_{min}, y^{i}_{min}) \mid x,y \in O^{i}\} \qquad (3.6)$$

$$B^{i}_{max} = \{(x^{i}_{min}, y^{i}_{min}) \mid x,y \in O^{i}\} \qquad (3.7)$$

Where O^i denote set of coordinates of points in the interest moving object I, $B^i_{max}$is the left top corner coordinates of the interest moving object i, and B^i_{max} is the right bottom corner coordinates of the interesting moving object i. In the chapter 4 shows the bounding box of the object tracking.Computerized following of items can be utilized by numerous fascinating applications.

A precise and effective following capacity at the heart of such a framework is fundamental for building larger amount vision-based knowledge. Tracking is not a trifling undertaking given the non-deterministic nature of the subjects, their movement, and the picture catch process itself. The goal of video tracking is to partner target objects in back to back video outlines. We need to identify and track the item moving freely to the foundation.there are many challenges in detecting of an object and tracking of objects and also recognition for fixed camera network .To set up a system for automatic segmentation and tracking of moving objects in stationary camera video scenes, which may serve as a foundation for higher level reasoning tasks and applications, and make significant improvements in commonly used algorithms. From the above discussion we found after tracking and detecting object we can extract its features for analysis of any undergone object by considering its track or motion by studying its feature points for building larger amount vision-based knowledge.

3.3.6Kalman Filter

A Kalman filter is used to estimate the state of a linear system where the state is assumed to the distributed by a Gaussian. The Kalman filter is a recursive predictive filter that is based on the use of state space techniques and recursive algorithms. It is estimated the state of a dynamic system. This dynamic system can be disturbed by some noise, mostly assumed as white noise. To improve the estimated state the Kalman filter uses measurements that are related to the state but disturbed as well. Kalman filtering is composed of two steps. Thus the Kalman filter consists of two steps:

- ˆ The prediction
- ˆ The correction

In the first step the state is predicted with the dynamic model. The prediction step uses the state model to predict the new state of the variables.

$\bar{X} = DX^{t-1}$ + W (3.8)

$\sum t$ = D $\sum t - 1$ D^t+ Q^t (3.9)

Where X^t and $\sum t$ are the state and covariance predictions at time t. D is the state transition matrix which defines the relation between the state variables at time t and t-1. Q is the covariance of the noise W. Similarly the correction step uses the current observation Z^t to update the object state

$$K^t = \sum t \, M^t [\, \mathrm{M} \textstyle\sum_m^t D^t + R^t] \qquad (3.10)$$

$$X^t = \mathrm{X} + K^t [Z^t - \mathrm{M} X^t] \qquad (3.11)$$

$$\sum t = \sum t - K^t \mathrm{M} \sum t \qquad (3.12)$$

where M is the measurement matrix, K is the Kalman gain which is called as the Riccati equation used for propagation of the state models. The updated state X^t is distributed by Gaussian. Similarly Kalman filter and extended Kalman filter assumes that the state is distributed by a Gaussian. In the second step it is corrected with the observation model, so that the error covariance of the estimator is minimized. In this sense it is an optimal estimator.

Kalman filter has been extensively used in the vision community for tracking there are many challenges in detecting of an object and tracking of objects and also recognition for fixed camera network. To set up a system for automatic segmentation and tracking of moving objects in stationary camera video scenes, which may serve as a foundation for higher level reasoning tasks and applications, and make significant improvements in commonly used algorithms. It particularly defines the initial and final positions of any object at its trajectories.

3.3.6.1 State Vector

The state vector contains the variables of interest. It describes the state of the dynamic system and represents its degrees of freedom. The variables in the state vector cannot be measured directly but they can be inferred from the values that are measurable. Elements of the state vector can be positioned, velocity, orientation angles, etc. A very simple example is a train that is driving with a constant velocity on a straight rail. In this case the train has two degrees of freedom, the distance and the velocity. The state vector has two values at the same time; one is the predicted value before the update and the posterior value after the update.

3.3.6.2 The Discrete Kalman Filter

In 1960, R.E. Kalman published his famous paper describing a recursive solution to the discrete data linear filtering problem. Since that time, due in large part to advances in digital computing; the Kalman filter has been the subject of extensive research and application, particularly in the area of autonomous or assisted navigation. The main problem with Kalman filtering is that statistical models are required for the system and the measurement instruments. Unfortunately, they are typically not available, or difficult to obtain. The two most commonly recommended methods of approaching this problem are:

- Employ an adaptive algorithm which adjusts these unknown parameters (such as the measurement noise variance) after each time step based on the observed measurements. This also accounts for processes with changing parameters.

- Perform an ”online” analysis of the system and measurement instruments prior to running the process (system identification). It should be noted however that the second approach will not always be applicable if the process cannot be observed directly. In other words, if the measurements in the online analysis also contain errors, the process cannot be accurately profiled.

Each $x_t$contains an m×1 mean vector ∧X and an m × n covariance matrix P, where m is the number of parameters that describe the state. A simple example of the parameters necessary for tracking are the x and y coordinates as well as the u and v velocity components. The yt nodes are represented by an n×1 vector which is nothing but the observed position of the target in the context of video tracking. This method of describing the state through a finite set of parameters is known as the state-space model (SSM).

As mentioned earlier, the state nodes are related to each other through the physics underlying object motion. The transition from one state to the next could be described in many ways. These different alternatives can be grouped into linear and nonlinear functions describing the state transition. Although it is possible to handle either of these transition types, the standard Kalman filter employs a linear transition function [2].

The Extended Kalman Filter (EKF) allows a non-linear transition, together with a non-linear measurement relationship. For the standard Kalman filter, the state transition from t to t + 1 can be expressed by the equation

$$X_{t+1} = \mathrm{A}x_t + W_t \quad (3.13)$$

Where A is referred to as the state transition matrix and w is a noise term. This noise term is a Gaussian random variable with zero mean and a covariance matrix Q, so its probability distribution is

$$p(w) \sim N\,(O,Q) \quad (3.14)$$

The covariance matrix Q will be referred to as the process noise covariance matrix in the remainder of this report. It accounts for possible changes in the process between t and t + 1 that are not already accounted for in the state transition matrix. Another assumed property of w is that it is independent of the statex_t. The measurement is taken, the node y becomes observed and x node can be

$$y_t = \mathrm{C}x_t + w_t \quad (3.15)$$

Where C is an m×n matrix which relates the state to the measurement. Much like wt. Now that a graphical model has been established and the relationships between the nodes are formulated, it is possible to look at how these relationships can be used in the tracking process. A prediction is required at each time step before the target is located with the tracking algorithm. The predicted position is nothing but the expected measurement given all the previous measurements. It illustrates the situation before each prediction is made and serves as a template for the recursive step. Initially Y_t is observed, and a prediction for $Y_{t+1}$is required. After the prediction is made and measurement is taken, y becomes observed, and the process t+1 repeats for t + 2 [2]. Tracking more than just the x and y coordinates of the target would be interesting and is possible. It would be especially interesting to study the applicability of the predication methods to these further degrees of freedom such as scaling or rotation.

When thinking along these lines, why not focus on a single dimension rather than using a 2D image and allowing both x and y translations? This approach was considered, and would include the analysis of one-dimensional intensity vectors rather than 2D images. It is in fact expected to provide clearer results. However, to study the effects of prediction on real image sequences, a 2D implementation was required. Furthermore, even though having results in an abstract setting is useful; a model with features a little more similar to real image sequences was desired. On average, the results achieved with the Kalman filter should be at least as good as those of the simple prediction method. This is expected to be helpful in the experiments to measure the performance of Kalman filter.

3.3.6.3 Discrete Kalman Filter Algorithm

The Kalman filter estimates a process by using a form of feedback control: the filter estimates the process state at some time and then obtains feedback in the form of (noisy) measurements. As such, the equations for the Kalman filter fall into two groups:

- Time update equations
- ˆ Measurement update equations

The time update equations are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the next time step. The measurement update equations are responsible for the feedback-i.e. for incorporating a new measurement into the improved estimate. The time update equations can also be thought of as predictor equations, while the measurement update equations can be thought of as corrector equations. The specific equations for the time and measurement updates are presented below

Discrete Kalman filter time update equations

$$X_k = \mathrm{A}x_{k-1} + \mathrm{B}u_{k-1} \tag{3.16}$$

$$P_k = \mathrm{A}P_{k-1_{A^T}} + \mathrm{Q} \tag{3.17}$$

Discrete Kalman filter measurements update equations

$$K_k = P_K^- H^T + P_K \mathrm{H} + R^- \qquad (3.18)$$

$$X_k = X_k^- - \mathrm{H}\hat{}x_k \qquad (3.19)$$

This recursive nature is one of the very appealing features of the Kalman filter it makes practical implementations much more feasible than an implementation of a which is designed to oper ate on all of the data directly for each estimate. The Kalman filter instead recursively conditions the current estimate on all of the past measurements.

3.3.6.4 Algorithm Discussion

The Kalman filter estimates a process by using a form of feedback control. The filter estimates the process state at some time and then obtains feedback in the form of measurements. As such, the equations for the Kalman filter fall into two groups: time update equations and measurement update equations. The time update equations are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the a priori estimates for the next time step. The measurement update equations are responsible for the feedback-i.e. for incorporating a new measurement into the a priori estimate to obtain an improved a posterior estimate. The time update equations can also be thought of as predictor equations, while the measurement update equations can be thought of as corrector equations. Indeed the final estimation algorithm resembles that of a predictor-corrector algorithm as shown. The specific equations for the time and measurement updates are presented below.

Discrete Kalman Filter Time Update

$$x_k^- = \mathrm{A}x_{K-1} +\_ \; \mathrm{B}u_{k-1} \qquad (3.20)$$

$$P_k^- = \mathrm{A}P_{K-1}A^T + \mathrm{Q} \qquad (3.21)$$

As shown the time update equations above project the state and covariance estimates forward from time step to step.

Discrete Kalman filter measurement update

$$K_k = P_k^- H^T \; (\mathrm{H}P_k^- H^T + \mathrm{R}) \qquad (3.22)$$

$$X_k = x_k^- + K_k \,(z_k - Hx_k^-) \qquad (3.23)$$

$$P_k = (\mathrm{I} - K_k \mathrm{H})P_k^{-1} \qquad (3.24)$$

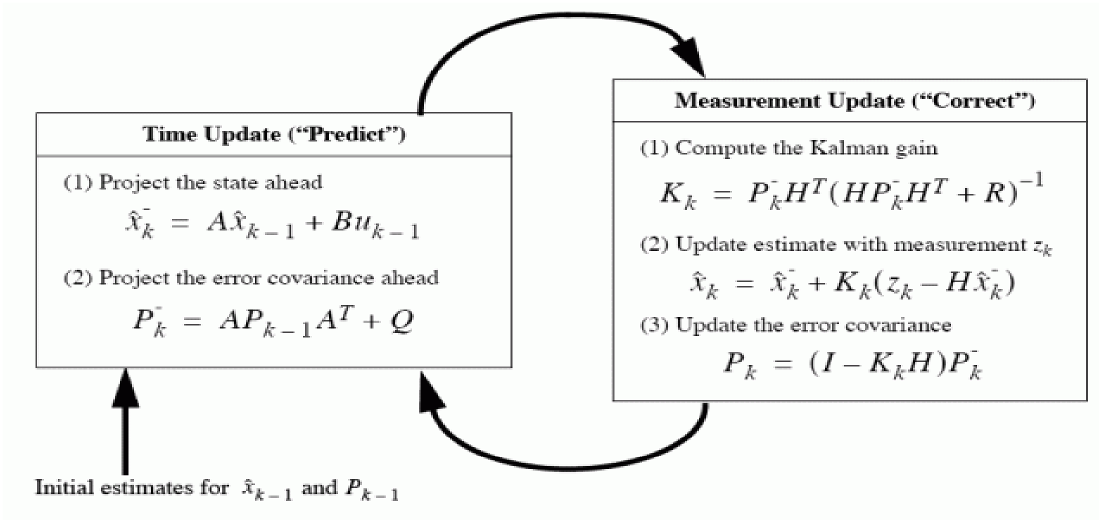


Figure 3.9: The discrete kalman Filter cycle [2]

As such, the equations for the Kalman filter fall into two groups: time update equations and measurement update equations. The time update equations are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the a priori estimates for the next time step. The measurement update equations are responsible for the feedback-i.e. for incorporating a new measurement into the a priori estimate to obtain an improved a posterior estimate. The time update equations can also be thought of as predictor equations, while the measurement update equations can be thought of as corrector equations. Indeed the final estimation algorithm resembles that of a predictor-corrector algorithm as show in figure. The specific equations for the time and measurement updates are presented below.

3.3.6.5 Filter Parameter And Tuning

In the actual implementation of the filter, the measurement noise covariance is usually measured prior to the operation of the filter. Measuring the measurement error covariance is generally practical (possible) because it is possible to measure the process anyway (while operating the filter) so generally it is possible to take

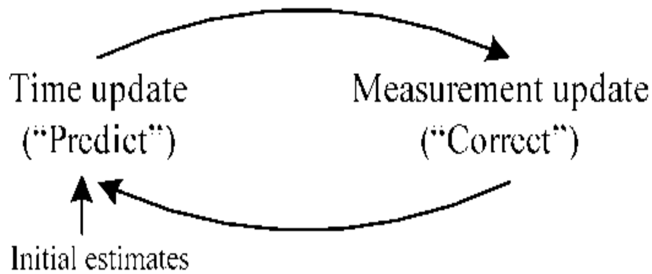


Figure 3.10: Operation Of kalman Filter cycle[34]

some off-line sample measurements in order to determine the variance of the measurement noise. The determination of the process noise covariance is generally more difficult as it not feasible to directly observe the process estimation. Sometimes a relatively simple process model can produce acceptable results if one "injects" enough uncertainty into the process via the selection of. Certainly in this case one would hope that the process measurements are more reliable. It is frequently the case that the measurement error is not constant and the process noise is also sometimes changes during filter operation. The main problem with Kalman filtering is that statistical models are required for the system and the measurement instruments. Unfortunately, they are typically not available, or difficult to obtain. The two most commonly recommended methods of approaching this problem are:

- Employ an adaptive algorithm which adjusts these unknown parameters (such as the measurement noise variance) after each time step based on the observed measurements. This also accounts for processes with changing parameters.
- Perform an online analysis of the system and measurement instruments prior to running the process (system identification). It should be noted however that the second approach will not always be applicable if the process cannot be observed directly. In other words, if the measurements in the online analysis also contain errors, the process cannot be accurately profiled.

3.3.6.6 Motion model

Motion models are a simple type of predictor and are quite common among simple systems. Motion models aims is to predict the next position based on a number of past observations. There may or may not make use of acceleration and can be expressed as

$$P(t+1) = P(t) + V(t) \qquad (3.25)$$

where P (t + 1) is the expected position at the next time step, P (t) is the position at current time step, and V (t) is the velocity at the current time step. For the simplest implementation,

$$V(t) = P(t) - P(t-1) \qquad (3.26)$$

3.3.7 Advantages of algorithm

- It is used for detection in thermal scenes.
- Tracking and visualization helps in position correction GPS system
- Implementation of completed model diagram can be used in surveillance.
- Object can be of small size or large size, so the algorithm can be adjusted.

3.3.8 Disadvantages of algorithm

- Slower speed of algorithm implementation.
- Large memory consumption reduces the implementation feasibility.
- Calculations performed using MATLAB is not feasible using other software's.
- Due to its video implementation it is hard to predict the actual video representation.

CHAPTER 4

STATEMENT FORMULATION AND PRESENTATION OF PROBLEM

4.1 Simulative result of detection and tracking algorithm

This chapter gives the idea about the existing and some modified methods analysis of algorithms for detection and tracking of objects. First some existing algorithm for detecting the objects like Frame difference method, Gaussian Mixture model to detect the object is discussed. Finally, background subtraction and background modeling is shown. After implementing all these existing algorithms then put one modified model for background modeling.

Tracking is the process of object of interest within a sequence of frames,from its first appearance to its last.The type of object and its description within the system depends on the application. During the time that it is present in the scene it may be occluded by other objects of interest or fixed obstacles within the scene.A tracking system should be able to predict the position of any occluded objects.

Object tracking systems are typically geared towards surveillance application where it is desired to monitor people or vehicles moving about an area. The ball tracking system has become a standard feature of tensile and cricket broadcast and uses object tracking techniques to locate and track the ball as it moves in the court. First implementation of an existing algorithm for tracking the object by using Block matching method is done.

4.2 Video Surveillance

An automatic video surveillance is used by private companies, governments and public organizations to fight against terrorism and crime, public safety in stairss, bus stand, railway station, town centers and hospitals. It has also find applications in traffic surveillance for efficient management of transport networks and road safety. Video surveillance system include task such as motion detection, tracking, and activity recognition. Out of the task mentioned above, detection of moving object is the first important step and successful segmentation of moving foreground object from the background ensures object classification, personal identification, tracking, and activity analysis, making these later step more efficient. Hu et al. [13] categorized motion detection into three major classes of method as frame differencing, background subtraction and Gaussian mixture.

4.3 Frame difference method

Frame differencing is a pixel-wise differencing between two or three consecutive frames in an image sequence to detect regions corresponding to moving object such as human and vehicles. The threshold function determines change and it depends on the speed of object motion. It's hard to maintain the quality of segmentation, if the speed of the object changes significantly. Frame differencing is very adaptive to dynamic environments, but very often holes are developed inside moving entities.Videos are actually consists of sequences of images, each of which called as a frame. For detecting moving objects in video surveillance system, use of frame difference technique from the difference between the current frame and a reference frame called as 'background image' is shown. That method is known as frame difference method. Frame differencing is the simplest moving object detection method which is based on determining the difference between input frame intensities and background model by using pixel per pixel subtraction. Grad. Sch. of Eng. et al. [5] have proposed frame difference method to detect the moving objects. In this case, frame difference method is performed on the three successive frames, which are between frame$|f_k - f_{k-1}|$and also the frame between$f_k$and $f_{k-1}$the output image as frame difference image is two difference images $d_{k-1}$and d_{k+1}.

$$d_{k-1}=|f_k - f_{k-1}| \quad (4.1)$$

$$d_{k+1} = |f_k - f_{k+1}| \quad (4.2)$$

$$d_{k'}(x,y) = \begin{cases} 1, & if\ d_{k'}(x,y) > T \\ 0, & otherwise \end{cases} \quad (4.3)$$

Where k' = k-1 and k+1

The process is followed by applying and operator to $d_{k-1}$and d_{k+1}This method is already discussed in details in chapter (2). Here original frames are shown and after preprocessing segmented results frames

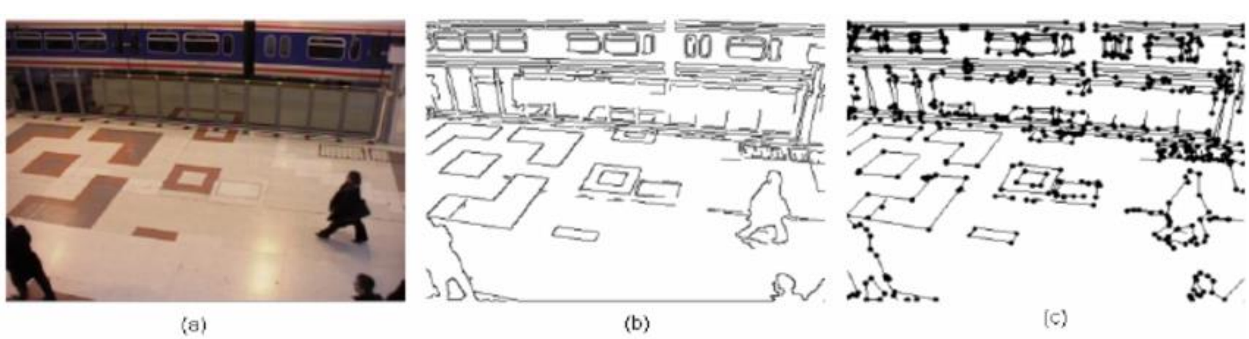


Figure 4.1: Original video frame images

4.4 Foreground detection

The foundation subtraction [10] is the most prominent and normal methodology for movement identification. The thought is to subtract the present picture from a reference foundation picture, which is upgraded amid a timeframe. It functions admirably just in the nearness of stationary cameras. The subtraction leaves just non-stationary or new protests, which incorporate whole outline area of an item .This methodology is basic and computationally affordable for ongoing frameworks, yet are to a great degree delicate to element scene changes from lightning and unessential occasion and so on. In this manner it is exceedingly reliant on a decent foundation upkeep model. Here in this part reproduction of different foundation subtraction procedures accessible in the writing, for movement division of article is performed. Foundation subtraction recognizes moving districts in a picture by taking the difference between the present picture and the reference foundation picture caught from a static foundation amid a timeframe. The subtraction leaves just non-stationary or new questions, which incorporate whole outline locale of an article. The issue with foundation subtraction [14], [8] is to consequently overhaul the foundation from the approaching video casing and it ought to have the capacity to defeat the accompanying issues: Motion out of sight: Non-stationary foundation districts, for example, branches and leaves of trees, a flag waving in the wind, ought to be identified as a component of the foundation.

Figure 4.2: output after frame difference

- Illumination changes: The background model should be able to adapt, to gradual changes in illumination over a period of time.
- Memory: The background module should not use much resource, in terms of computing power and memory.
- Shadows: Shadows cast by moving object should be identified as part of the background and not foreground.
- Camouflage: Moving object should be detected even if pixel characteristic are similar to those of the background
- Bootstrapping: The background model should be able to maintain background even in the absence of training background (absence of foreground object).

4.5 Simple Foreground detection

In simple background subtraction a absolute difference is taken between every current image I_t(x,y) and the reference background image B(x,y) to find out the motion detection mask D(x,y). The reference background image is generally the first frame of a video, without containing foreground object.

$$D(x,y) = \begin{cases} 1 & if\ |I_t\,(x,y) - B(x,y)| \geq \tau \\ 0 & otherwise \end{cases} \qquad (4.4)$$

where τ is a threshold, which decides whether the pixel is foreground or background. If the absolute difference is greater than or equal to τ , the pixel is classified as foreground, otherwise the pixel is classified as background.

4.6 Running Average

Simple background subtraction cannot handle illumination variation and results in noise in the motion detection mask. The problem of noise can be overcome, if the background is made adaptive to temporal changes and updated in every frame.

4.7 Morphological Operation

Morphological operations apply an organizing component to an information picture, making a yield picture of the same size. Morphological operation is performed to fill little crevices inside the moving protest and to decrease the clamor stayed in the moving items. The morphological administrators executed are widening trailed by disintegration .In expansion, every foundation pixel that is touching an article pixel is changed into an item pixel. Expansion adds pixels to the limit of the article and closes secluded foundation pixel. Expansion of set A by organizing component B [7] is defined as :

$$A \oplus B = \cup_{b \in B} (A)_b \qquad (4.5)$$

In erosion, each object pixel that is touching a background pixel is changed into background pixel. Erosion removes isolated foreground pixels. Erosion of set A by structuring element B [7] is defined as:

$$A \ominus B \quad = \quad \cup_{b \in B} (A)_{-b} \qquad (4.6)$$

The number of pixels added or removed from the objects in an image depends on the size and shape of the structuring element used to process the image. Morphological operation eliminates background noise and fills small gaps inside an object. There is no fixed limit on the number of times dilation and erosion is performed. In the given algorithm dilation and erosion is used iteratively till the foreground object is completely segmented from the background. After morphological operation now the results of following frames, remove noise from frame difference and background subtraction frame result.

4.8 Gaussian Mixture Model

To actualize a current Gaussian mixture model in light of foundation model to distinguish the moving items. For distinguishing moving items in video observation framework the utilization the Gaussian mixture model, is fundamental this model has the shading estimations of a specific pixel as a blend of Gaussians. In any case, the pixel values that don't fit the foundation disseminations are considered as closer view. Nowak 2003 demonstrated how the parameters of a blend of Gaussians for which every hub of a sensor system had different blending coefficients could be evaluated utilizing a conveyed adaptation of the understood desire boost (EM) calculation. This message-passing calculation includes the transmission of sufficient insights between neighboring hubs in a specific arrange, and was tentatively appeared to focalize to the same results as concentrated EM. Kowalczyk and Vlas-sister Kowalczyk and Vlassis, 2004 [8] proposed a related tattle based disseminated calculation called Newscast EM for assessing the parameters of a Gaussian blend. Irregular sets of hubs over and again trade their parameter gauges and consolidate them by weighted averaging. In this area, another strategy that is usually utilized for performing foundation division. Stauffer and Grimson et al. [5]have proposed; recommend a probabilistic methodology utilizing a blend of Gaussian for recognizing the foundation and forefront objects. The likelihood of watching a given pixel esteem Pt at time t is given by.

$$P(p_t) = \sum_{i=1}^{k} w_{i,t}\, (p_{t,} \mu_{i,t,} \sum i, t) \qquad (4.7)$$

Where k is the number of Gaussian Mixture and that is used. The number of k varies depending on the memory allocated for simulations. Then the normalized Gaussian η is a function of $w_{i,t}\, \mu_{i,t,} \sum i, t$ which represents weight, mean and co-variance matrix of the ith Gaussian at time respectively.The weight indicates the influence of the ith Gaussian and time t. In this case k=5 to maximize the distinction amongst pixel values. Since it is an iterative process that all parameters are updated, with the inclusion of every new pixel. Before update take place, then the new pixel is compared to see if it matches any of the k existing Gaussian. A match is determined if $|p_{t,} - \mu_{i,t,}| < 2.5\sigma$

Where correspond to the standard deviation of the Gaussian.For distinguishing moving items in video observation framework the utilization the Gaussian blend model, is fundamental this model has the shading estimations of a specific pixel as a blend of Gaussians. In any case, the pixel values that don't fit the foundation disseminations are considered as closer view Depending on the match, the Gaussian mixture is updated in the following manner:

$w_{i,t}$= (1-α)$w_{i,t-1}$+α (4.8)

$\mu_{i,t}$ = (1-ρ)μ_{t-1}+$p_t\rho$ (4.9)

$\sigma^2_{i,t}$= (1-ρ)$\sigma^2_{i,t-1}$+ ρ(p_t- μ_t)T (p_t- μ_t) (4.10)

Where $\rho = \ \alpha\eta((\rho_t|\mu_{i,t-1,}\sigma_{i,t-1})$ (4.11)

In this case the variable (1|α) defines the speed at which the distribution parameter changes. In the pixel (pt) matches the i-th Gaussian, then the matching remaining (k-1) Gaussians are updated in the following manner,

$w_{i,t}$= (1-α)$w_{i,t-1}$ (4.12)

$\mu_{i,t}$ = $\mu_{i,t-1}$ (4.13)

$\sigma^2_{i,t}$ = $\sigma^2_{i,t-1}$ (4.14)

The values for weight and variance vary based on the significance that is given to a pixel which is least likely to occur in a particular way. All the Gaussian weights are normalized after the update is performed. The k-Gaussians are then reordered based on their likelihood of existence.Then (b) distribution are modeled to be the background and the remaining (k-b) distributions are modeled as the foreground for the next pixel.

The values for (b) is determined

B = $argmin_b\left(\sum_{i=1}^{b} w_i \ > T\right)$(4.15)

Where T is some threshold value which measures the proportion of the data that needs to match the background and then the first B distribution are chosen as background model.

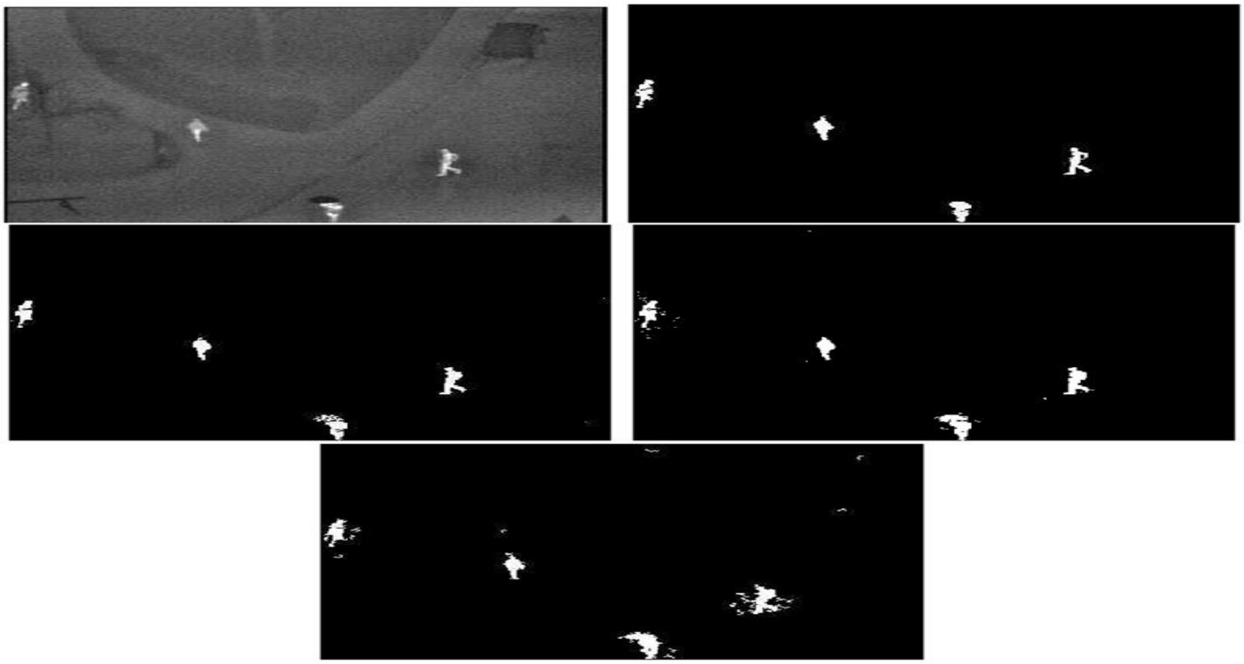


Figure 4.3: Gaussian mixture

4.9W^4 foreground detection

W^4model is a simple and effective method for segmentation of foreground objects from video frame. In the training period each pixel uses three values; minimum m(x,y), maximum n(x,y) and the maximum intensity difference of pixels in the consecutive frames d(x,y) for modeling of the background scene. The initial background for a pixel location (x, y) is given by [15]

$$\begin{matrix} m(x,y) & & min_x V^z(x,y) & \\ n(x,y) & = & max_x V^x(x,y) & where'z'are\ the\ frames\ satisfying. \\ d(x,y) & & max_x |v^z(x,y) - V^{Z-1}(x,y)| & \end{matrix}$$

The background cannot remain same for a long period of time, so the initial back-ground needs to be updated. W^4uses pixel-based update and object based update method to cope with illumination variation and physical deposition of object. W^4uses change map for background updation.

A detection support map (gS) computes the number of times the pixel (x, y) is classified as background pixel. A detection support map gS computes the number of times the pixel (x, y) is classified as background pixel.

$$gS_t(x,y) = \begin{cases} gS_{t-1}(x,y) + 1 & if\ pixel\ is\ background \\ gS_{t-1}(x,y) & if\ pixel\ is\ foreground \end{cases} \quad (4.16)$$

A motion support map (mS) computes the number of times the pixel(x, y) isclassified as moving pixel.

$$\mathrm{m}S_t(\mathrm{x,y}) = \begin{cases} mS_{t-1}(x,y)+1 & if\ M_t(x,y)=1 \\ mS_{t-1}(x,y) & if\ M_t(x,y)=0 \end{cases} \quad (4.17)$$

where

$$M_t(\mathrm{x,y}) = \begin{cases} 1\ if\ (|I_t(x,y)-I_{t+1}(x,y)|>2*\sigma)\wedge \\ (|I_{t-1}(x,y)-I_t(x,y)|>2*\sigma) \\ 0\ otherwise \end{cases} \quad (4.18)$$

4.10 Foreground Modeling

The fundamental rule of foundation subtraction is to contrast a static foundation outline and the present edge of the video scene pixel by pixel. This system constructs a model of the foundation any edge can be contrasted with the model with recognize zones where a significant difference happens. A few stages for the foundation subtraction, i.e. first advancement of a foundation model of the scene is done, then foundation subtraction to distinguish forefront object. Julio cezar et al. [3] has proposed a foundation demonstrate, and joins a novel strategy for shadow identification in dark scale video groupings. In the first stage, a pixel insightful middle filter after some time is connected to a few seconds of recordings (commonly 20-40 seconds) to recognize moving pixels from stationary pixels.In the second stage, just those stationary pixels are handled to build the underlying foundation model. Give V a chance to be a cluster containing N sequential images,V k(i,j) be the force of a pixel (i,j) in the K-th pictures of V,σ(i,j) and λ(i,j) bethe standard deviation and middle estimation of intensities at pixel (i,j) in all pictures in V, individually. The underlying foundation for a pixel (i,j) is performed by a three-dimensional vector: the base m(i,j) and most extreme n(i,j) power values and the greatest force difference d(i,j) between the successive edges saw amid this preparation period. At that point the foundation model B(i,j)=[m(i,j), n(i,j), d(i,j)] is gotten As takes after: [m(i,j) n(i,j) d(i,j)] is acquired. As follows :

$$\begin{bmatrix} m(i,j) \\ n(i,j) \\ d(i,j) \end{bmatrix} = \begin{bmatrix} min_x V^z(i,j) \\ max_x V^x(i,j) \\ max_x |v^z(i,j) - V^{Z-1}(i,j)| \end{bmatrix}$$

Where 'z' are the frames satisfying $|V^z$ (i,j)λ(i,j)|≤ 2σ(i,j) After the training period, an initial background model B(i,j) is then obtained.

Then each input imageI_t(i,j) of the video sequence is compared to B(i,j), and a pixel (i,j) is classified as a foreground pixel if I_t(i,j) > (m(i,j) − kμ) and I^{ts}(i,j) < (n(i,j) + kμ)

After the background subtraction by using the above method, some shadow region pixels which are wrongly classified as foreground objects.so they have employed a shadow region detection method to remove the shadow pixels from detected foreground pixels.

4.10.1 Analysis

Here examination of different recognition calculation is finished. In the examination of taking different video, it is demonstrated that, Julio cezar technique is the best question recognition strategy. In the wake of getting the outline of the foundation protests, the shape of the forefront objects after some post preparing of the subsequent outline like locale filling is performed. At that point extraction of the component from the shape is finished. Be that as it may, the element extraction part is not talked about here**.**

4.10.2 Object Tracking

After the object detection is achieved, the problem of establishing a correspondence between object masks in consecutive frames should arise. Obtaining the correct track information is crucial for subsequent actions, such as object identification and activity recognition. For this situation, block matching technique is used.

4.10.3 Block Matching Method

The whole procedure of following the moving item is represented in the accompanying Fig 2.6. The square coordinating strategy is all around portrayed in [4], which we have connected here. Square coordinating is a procedure for following the enthusiasm moving item among the moving articles rising in the scene. In this article, the pieces are defined by separating the picture outline into non-covering square parts. The squares are made in light of fringe augmentation sign relationship (PISC) picture Satoh et al., 2001; Sugandi et al., 2007 [10] that considers the shine change in every one of the pixels of the pieces in respect to the considered pixel. Fig. 9 demonstrates the square in PISC picture with piece size is 55 pixels. In this way, one square comprises of 25 pixels. The pieces of the PISC picture in the past casing are defined as appeared in Eq. (2.3). Additionally, the squares of the PISC picture in the present edge are defined in Eq. (2.4). To decide the coordinating criteria of the pieces in two progressive edges, assessment is done utilizing relationship esteem that communicates as a part of Eq. (2.5). This mathematical statement computes the relationship esteem between piece in the past casing and the present one for all pixels in the square. The high relationship esteem demonstrates that the pieces are coordinated each other. The enthusiasm moving item is resolved when the quantity of coordinating pieces in the past and current edge are higher than the specific limit esteem. The edge worth is acquired tentatively.This system constructs a model of the foundation any edge can be contrasted with the model with recognize zones where a significant difference happens. A few stages for the foundation subtraction, i.e. first advancement of a foundation model of the scene is done, then foundation subtraction to distinguish forefront object. Julio cezar et al. [3] has proposed a foundation demonstrate, and joins a novel strategy for shadow identification in dark scale video groupings.

$$b_{np} = \begin{cases} 1, & if\ f_{np \geq f(i,j)} \\ 0, & otherwise \end{cases} \qquad (4.19)$$

$$b'_{np} = \begin{cases} 1, & if\ f_{np \geq f(i,j)} \\ 0, & otherwise \end{cases} \qquad (4.20)$$

$$corr^{n} = \sum_{P=0}^{N} b_{np} * b'_{np} + \sum_{P=0}^{N} (1 - b_{np}) * (1 - b'_{np}) \qquad (4.21)$$

Where b and b' are the blocks in the previous and current frame, n is the number of block and N is the number of pixels of block.

4.10.4 Motive of tracking

The tracking used in this article can be described as following. Firstly, blocks and the tracking area are made only in the area of moving object to reduce the processing time. The previous frame is divided into block size (block A) with 9x9 pixels in the previous frame. It is assume that the object coming firstly will be tracked as the interest moving object. The block A will search the matching block in each block of the current frame by using correlation value. In the current frame, the interest moving object is tracked when the object has maximum number of matching blocks. When that matching criteria is not satisfied, the matching process is repeated by enlarging the tracking area (the rectangle with dash line).The blocks still are made inside the area of moving object. When the interest moving object still cannot be tracked, then the moving object is categorized as not interest moving object or another object and the tracking process is begun again from the begin.

4.10.5 Feature Extraction

The element of items removed in the spatial space is the position of the followed object. The spatial data joined with the components in time area speaks to the direction of the followed object, so the development and pace of the moving articles can be evaluated that requirements to followed. Along these lines, the elements of spatial area are essential to protest identification. The jumping encloses defined Eq. (2.4) is utilized as spatial data of moving articles. In the wake of getting the enthusiasm moving article, then extraction of enthusiasm moving item by utilizing a bouncing box. The bouncing box can be dictated by registering the most extreme and least estimation of x and y directions of the enthusiasm moving item as per the accompanying mathematical statement:

whereOi indicates set of direction of focuses in the enthusiasm moving item i,Bimin is the left top corner coordinates of the enthusiasm moving article i, and Bimax is the high base corner coordinates of the intriguing moving article i.In the section 4 demonstrate the bouncing box of the item following.

- **Low level feature extraction**
  Low level feature (e.g., shading, angle, movement) presents a graph of shading representation, incline and development computational methods. The goal is to perceive how to try low level trademark in the differing period of target representation and impediment. Depicting shade through trademark that proceed as before paying little regard to moving circumstances included by picture taking care of is a basic need for some component taking after demand. Changes in imaging condition are related to the audit course, the objective surface presentation what's more, lighting up condition. This change present antiques, for instance, shading shadows and highlights. Close-by power changes convey vital data information about the appearance of object of speculation. This change happen inside the thing itself and the cutoff points amongst article and mapping the photo onto low level trademark may not be useful to achieve.
- **Mid-level feature extraction**
  Mapping a target summation of picture substance, thus lessening the adequacy of a feature tracker. So we break down feature using subset of pixel that address essential structure (e.g., edge, corner, interest point) or uniform district where Whole pixel impart some customary properties. Most speculation point identifier tends to choose exceedingly extraordinary neighborhood case, for illustration, corner, prominent edge and area with isolate surface establishment.
- **High level feature extraction**
  To describe an interested object, one could be aggregate mid-level characteristics, for instance, premium point and area or can perceive clearly the thing general centered on its appearance. Abnormal state trademark may be centroid, whole range or presentation of interested objects. A verity of procedures for distinguishing moving objects are engaged around background subtraction is known as background modeling.

4.11 Motion-Based Multiple Object Tracking

Identification of moving questions and movement based following are imperative segments of numerous PC vision applications, including action acknowledgment, traffic observing and car security. The issue of movement based article following can be gap into two sections Detecting moving items in every edge Associating the discoveries relating to the same article after some time The identification of moving items utilizes a foundation subtraction calculation in view of Gaussian blend models. Morphological operations are connected to the subsequent forefront veil to kill commotion. At last, blob investigation distinguishes gatherings of associated pixels, which are prone to compare to moving objects.The relationship of identifications to the same article is constructing exclusively in light of movement. The movement of every track is evaluated by a Kalman filter. The filter is utilized to anticipate the track's area in every casing, and decide the probability of every location being doled out to every track. In any given casing, a few recognitions might be allocated to tracks, while different identifications and tracks may stay unassigned. The allotted tracks are redesigned utilizing the relating identifications. The unassigned tracks are checked undetectable. An unassigned discovery starts another track. In movement recognition how numerous articles are followed by utilizing Kalmanfilter. At first framework items are made for perusing the video outlines. At that point kalman filter articles are utilized for movement based tracking.Then the aggregate number of edges in which tracks are detected.Here it is appeared, how movement based various objects are followed. The calculation includes two stages:

- Step 1: Compute the expense of allotting each discovery to every track utilizing the separation strategy. The cost considers the Euclidean separation between the anticipated centroid of the track and the centroid of the location. It additionally incorporates the confidence of the expectation, which is kept up by the Kalman filter.

- Step 2: Solve the task issue spoke to by the cost lattice utilizing the relegate Detections To Tracks capacity. The capacity takes the cost framework and the expense of not allocating any recognitions to a track.

The worth for the expense of not allocating recognition to a track relies on upon the scope of qualities returned by the separation strategy for the Kalman Filter. This quality must be tuned tentatively. Setting it too low improves the probability of making another track, and may bring about track fracture. Setting it too high may bring about a solitary track relating to a progression of independent moving items.

CHAPTER 5

RESULTS AND SIMULATION

5.1 Investigation and Discussion

The thesis aims to investigate the different scenarios in which detection and then tracking. The first and foremost results are evaluated on a basic video scheme that is depicting a simple scene. We have calculated for five videos in thesis and b calculated for more accordingly this all videos.

1. First video is about a simple stair scene in which a unknown baggage is put in an abandoned region in which the complete possibilities of surveillance could be explored without going much into the respective details, as the possible understanding is made on the basis of simple foreground so detected.

2. Second video was of real time in which every passenger came for checking and scanning their luggage and through x ray machine we can check luggage's and identify abandoned objects or suspicious objects.

3. Third video was of station where one person comes and left his one bag at his seat and after sometime unattended then through kalman filter we calculated its tracking and threshold values.

4. Fourth video is also an unattended luggage at some isolated place he left his luggage and through kalman filter we tracked it and detected.

5. Fifth video at a plane background is taken and then a ball is moving in the scene. At a time the ball is occluded and due to which the detection of ball and its tracking in its due course becomes relatively infeasible.

This thesis shows the possibilities if detecting the locations of unattended luggage's and object or a ball in occluded scenes in which the positions of these objects before occlusions are taken and then a estimate of that position is given as a primary input in kalman filter which is the main sole of this thesis. Let's start analysis with all these videos and calculating their processing time through various graphs.

Investigation Scene 1 : occluded position of bag

Figure 5.1: The scene of a stairs with passenger in the foreground highlighted with white

The results in the figure 5.6 is evalauted on this image for different people in the scenes that aims to reflect the motion of passengers. However this figure is aimed towards detecting and tracking an object as it will remain unchecked in the scene after few frames. This is reflected in the next frames available. The figure 5.8 reflects such scenes in black and white. This figure is shown in which the foreground object and people are all reflected as white. Due to high degree of motion in the figure the image seems to be good in terms of foreground detection and then the tracking by kalman will also be good which will be seen in the next few figures.

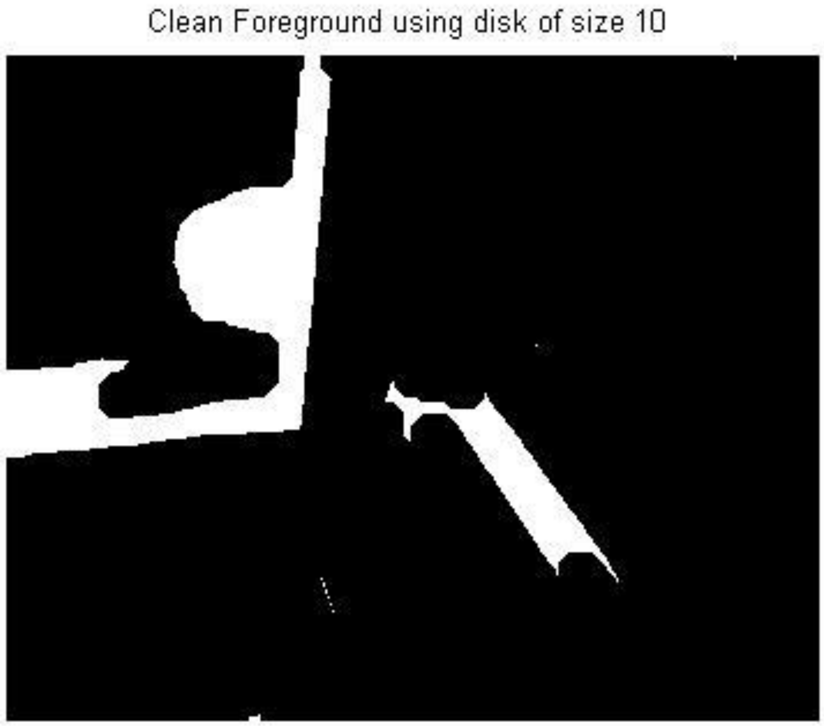


Figure 5.2: Noisy image at stairs

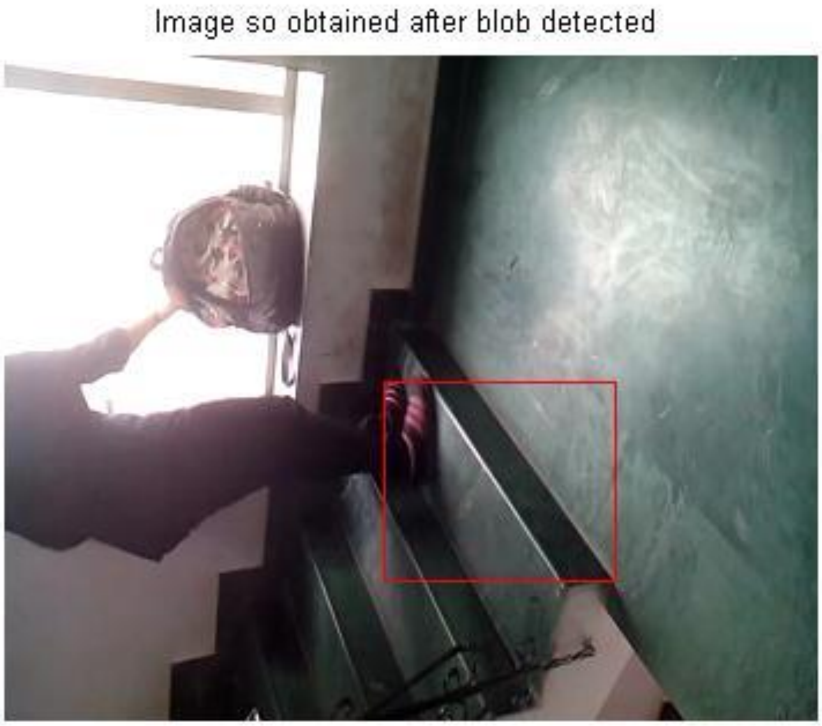


Figure 5.3 : This figure is the collection of accumulation of points of interests that are available in the detection.

The figure is a complete scene overlapping and description of points that are not good in the points form as well. This reflects a lot of occlusions that are present in the moving video. However the Gaussian mixture model available in this scene is doing good work and predicting the moving occlusions. In the next figure the occlusions and original points are all plotted together to show a good point combination of moving points. Several experiments had been done to evaluate the tracking algorithms. The sequences used in experiments consist of indoors and outdoors testing environment so that the proposed scheme can be fully evaluated. The target object is tracked in the next frame by individually comparing each feature point found from the next frame to those on the target object. The Euclidean distance is worked out. The candidate can be preserved when the two features Euclidean distance is larger than a threshold. So the good matches picked out by the consistency of their location, orientation and scale.

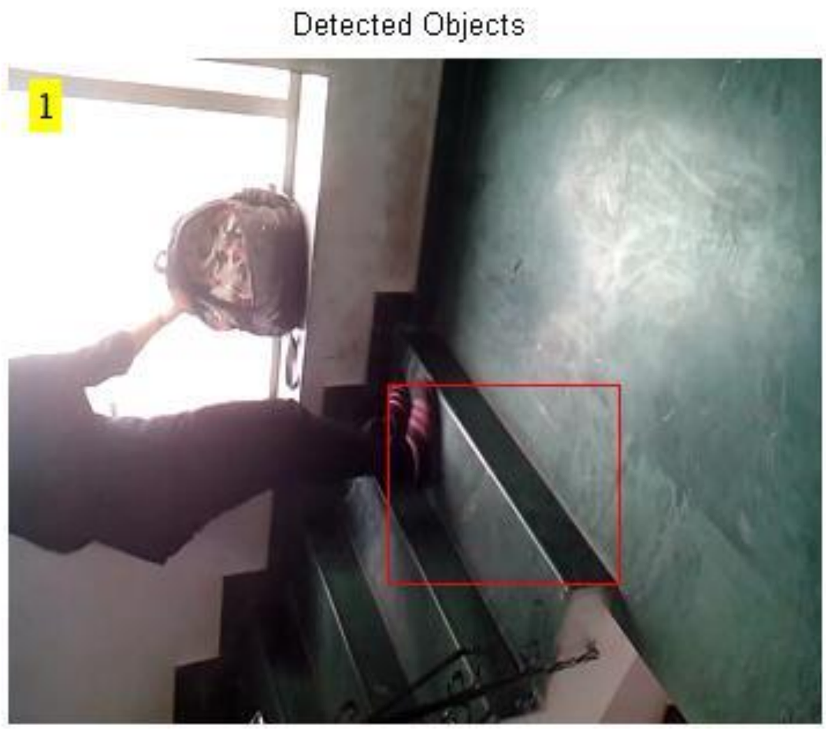


Figure 5.4: The figure shows a good amount of detection and tracking.

The existing figure shows a good amount of connection between the moving targets that are people and the actual people that are reflected in the occlusions. These degrees of collateral effects are due to non-linearity in the actual scene and moving scenes. Now the estimation made by kalman filter is upto the mark with some non-fictional points.

3. Investigation of scene: Real time at Stairs

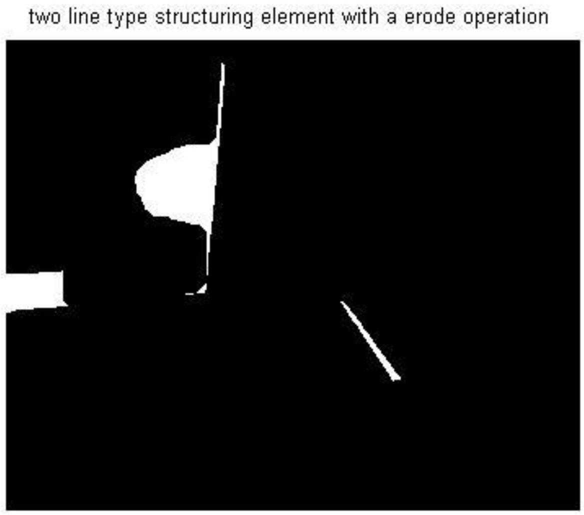


Figure 5.5 This figure is stairs checking survilence

In this figure 5.13 we can see number passengers are coming at stairs before going forward they have to check their luggage through multiple processing techniques or steps .At stairs security system and hence ,in this figure we can they have to keep luggage in tray provided and also any extra object they have then they are further processed for detecting ans tracking ,if or whether it contains any explosives or not then the person who drop there luggage for scanning n processing,they come there CT scanning of there body thrn after they collect there luggage.And tese steps follow for each and every passengers arriving at stairs.

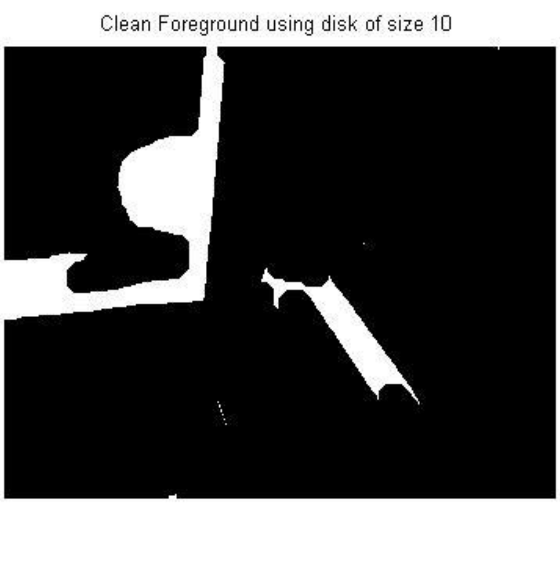


Figure 5.6: Background subtraction

In 5.6 figure background is subtracted for clear view of each and every passengers and there luggages at the stairs. This scene shows a high level of accuracy mentioned by Gaussian mixture model that is used in the existing code to modify the foreground. In the next few figures the results are seen in the form of tracking and detection by kalman filters.

Figure 5.7: end points detections

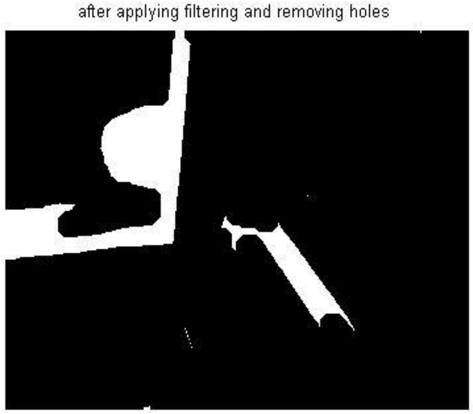


Figure 5.8: Analysis through kalman filter

In figure 5.9 and 5.10 is demonstration of detection of people done by Gaussian mixture model that helps in identifying the positions for estimation and tracking of people in the background scene. The kalman filter image which is figure 5.16 shows the connection of this people moving tighter in the scenes and occluded.

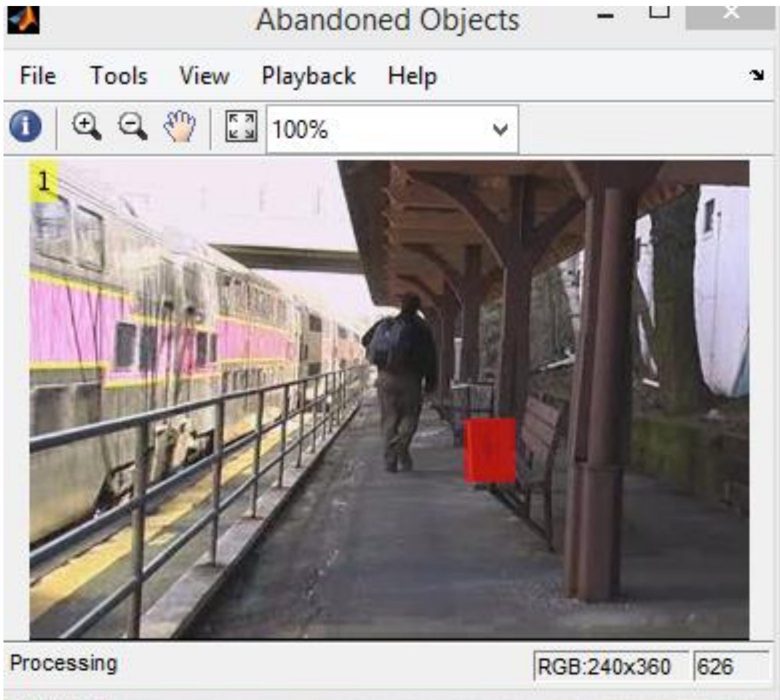


Figure 5.9: Display of abandoned object on a railway platform

Here the analysis has been extended to a railway platform where occurrences of missed objects, default objects , hidden objects and unknown objects are more frequent than in other cases. So the results which are shown in this scenario is of utmost importance that such codes in the existing research can detect and label those objects with a good level of timing and accuracy.

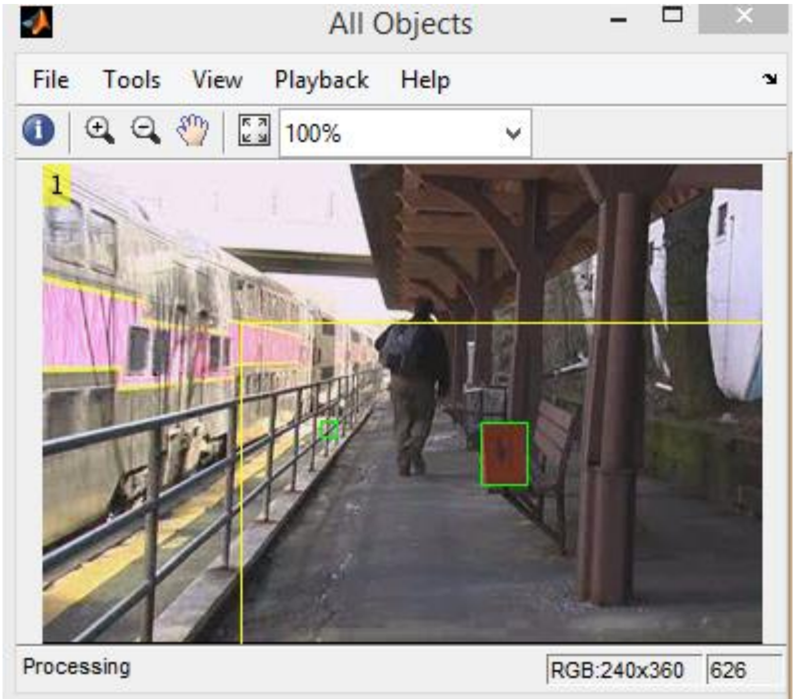


Figure 5.10 Region of interests labeling on the railway platform

This example snap shot added after figure 5.16 has been introduced to demonstrate a linking of object detection with region of interests. In this example a region of interests has been labeled at the earliest to show case that the complete object detected has been surrounded by a roi which reduces the complexity of entire searching mechanism in the full video length, thus helping in other user or system to locate the region and object easily.

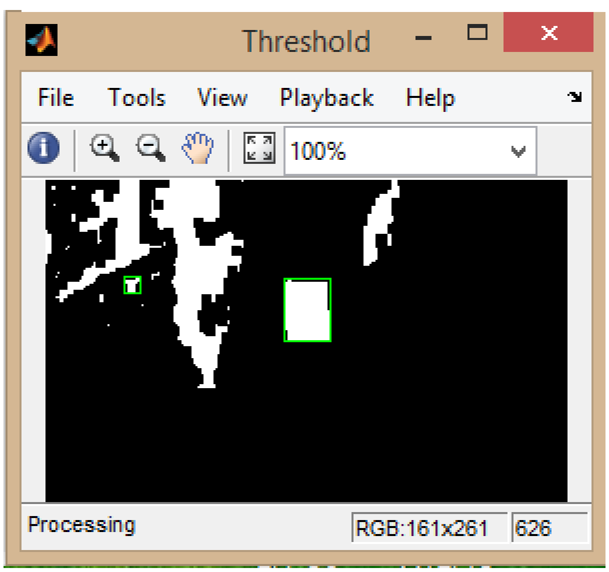


Figure 5.11.Binary display of abandoned object detection in the complete video of platform.

This figure indicates the blob analysis done in the black and white image as a result of the region of interests fulfilling the criteria of minimum blob area and its centroid location. This part of analysis is important before the actual detection is made, reason for such part is morphology is easily applicable to binaries white component and selection of structuring element to such white objects are easily configurable.

4..Investigation scene- Luggage scene

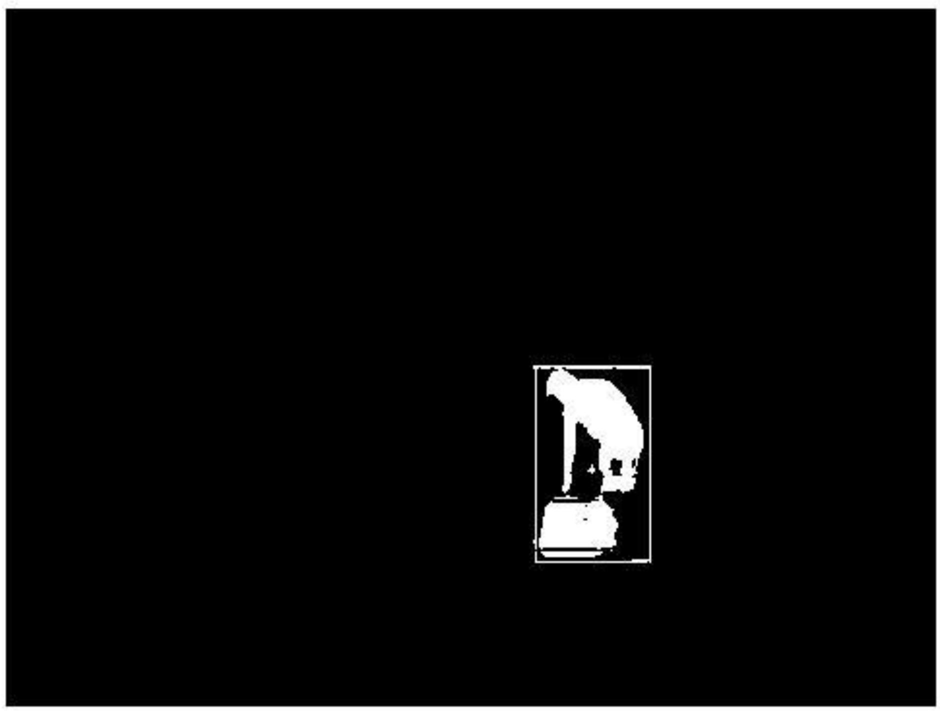
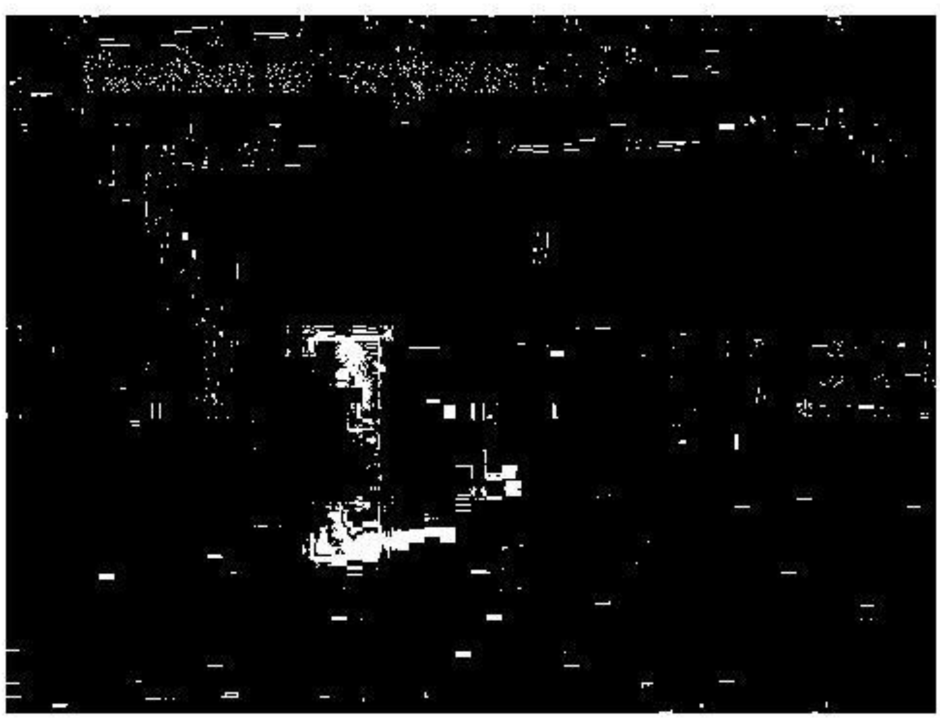


Figure 5.12 The figure shows a black and white effect of scene with obstacle detection .

This figure is a scene of a person showcasing a bag in its left and the complete scene is captured using a bounding box. This scene shows a high level of accuracy mentioned by Gaussian mixture model that is used in the existing code to modify the foreground. In the next few figures the results are seen in the form of tracking and detection by kalman filters.Figure 5.10 :

This figure is a example of noise introduced deliberately.The existing figure is demonstration of noise model introduced in the example to showcase the effect of non linearity that causeproblem in finding the background. Next figures are examples of tacking done by kalman filters.



Figures 5.13 The figure is a complete overlapped scenes that is a part of detection only.

The existing figure 5.13 is demonstration of detection of people done by Gaussian mixture model that helps in identifying the positions for estimation and tracking of people in the background scene. The kalman filter image which is figure 5.12 shows the connection of this people moving tighter in the scenes and occluded.

Noisy image

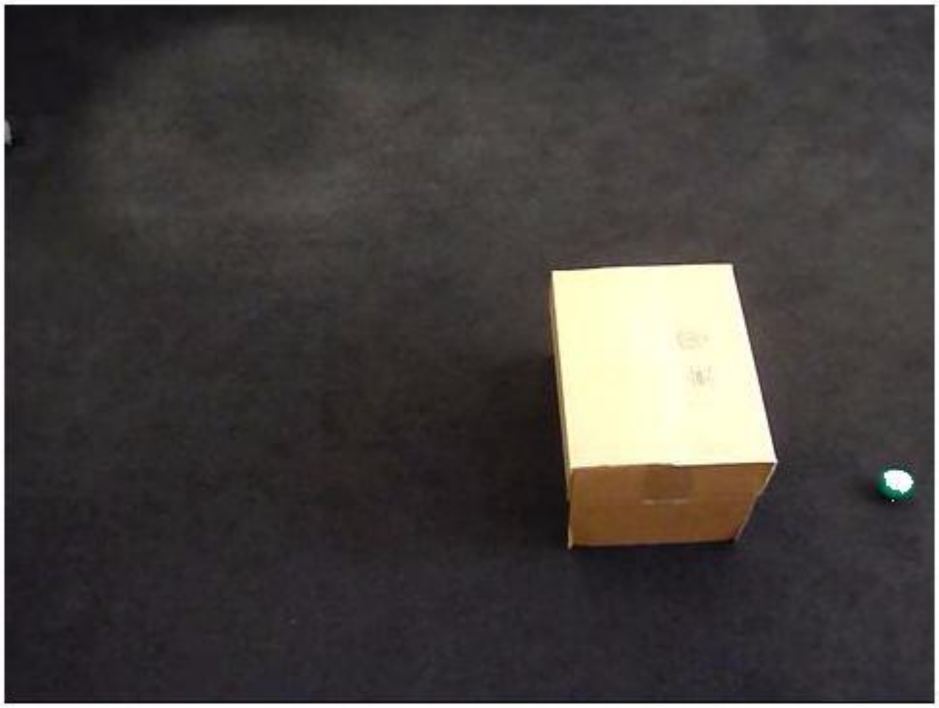


Figure 5.14 : The ball at one end of the scene

Now the ball background must be evaluated that has a good scene of results in which a black background and a white foreground depicting the results of noisy images.

Figure 5.15: This is the black and white foreground of the original image scene

This scene in black and white has a good theme of evaluating the background but this image is a result of Gaussian mixture model, which evaluates the foreground or background. Now as the coding section uploads the video the occlusion and its relational tracking in the image is found using the kalman filter whose results are evaluated as shown in figure 5.14 and figure 5.15.

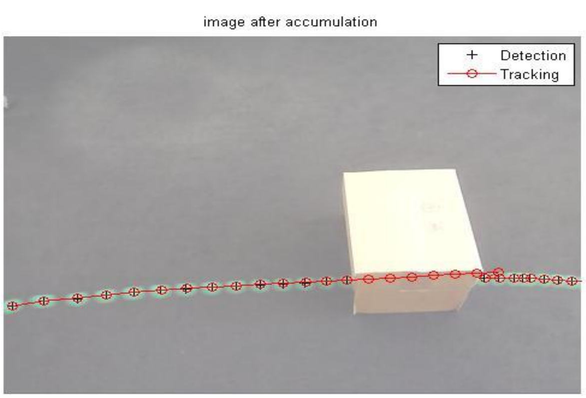


Figure 5.16 The results are found using detection in the scene for the positions of ball at different points using centroid

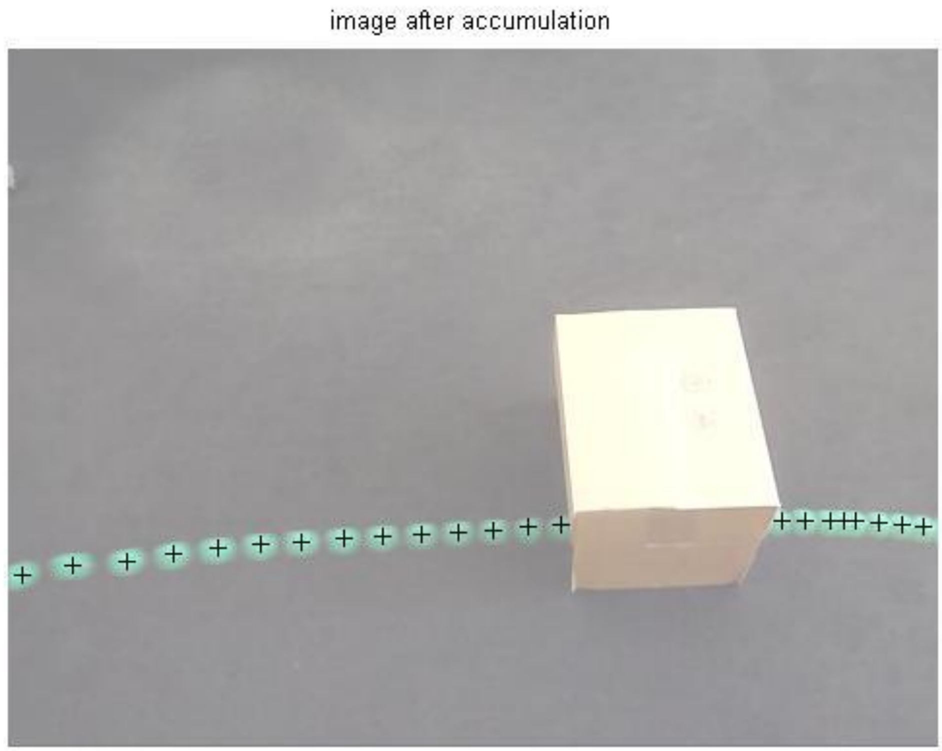


Figure 5.17 Image points predicted after applying first kalman filter initial position estimated values.

As a user and individual the points predicted by kalman filter in the figure 5.14 has a significant set of points showcasing the actual movement, but it is necessary to evaluate the points at those scenes where position of balls are occluded. Though this seems to be impractical the results are good in terms of estimated positions of the scenes using kalman filter. The results are evaluated and shown in figure 5.5 and 5.6 by applying kalman filter twice on the same scene. The two figure displays error to a margin level and then correcting.

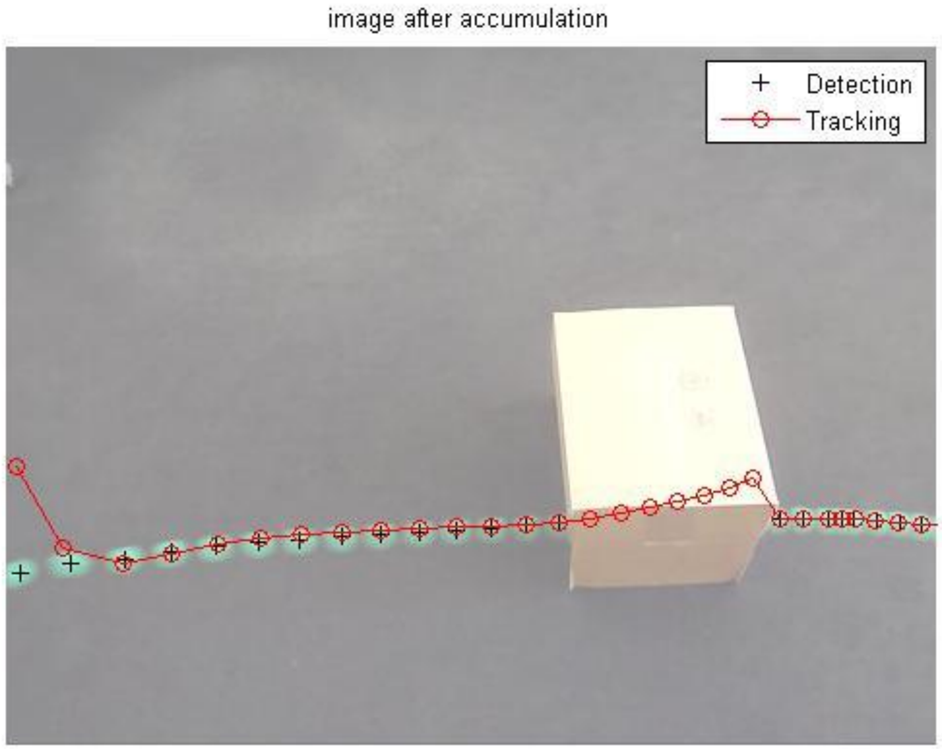


Figure 5.18 This is the first estimate of kalman filter in which positions are not upto the mark

Results evaluated for ball for detection and tracking after final estimation by kalman filter. This is the final investigation of kalman filter done in this thesis. Before the complete investigation of results could be complied with the predicted behavior it is necessary for the research work to conclude with new results which are also without any kind of implementation flaws.

In a similar manner the performance of these videos through kalman filter comes to graphs by writing their minimum and maximum time and also their processing time for covering up video timing and with significant specifications are calculated in different scenes, these graphs for different scenes are:

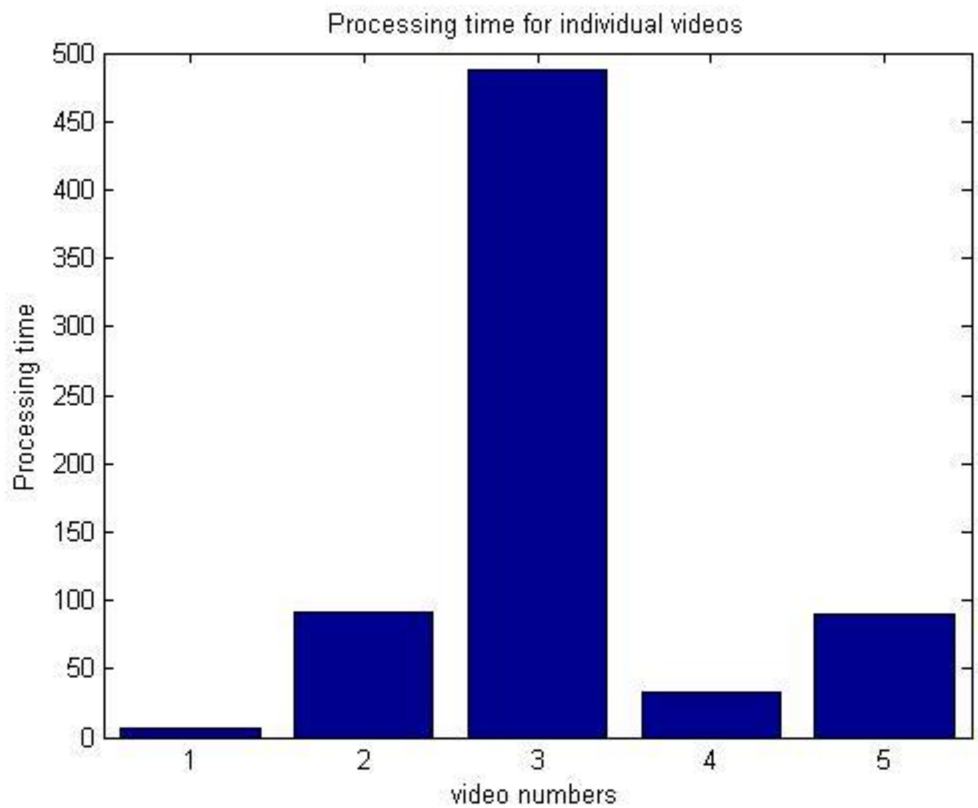


Figure 5.19: Processing time of above videos

This figure represents of five videos processing time. Above graph analysis in this graph figure shows that video three takes time approx. 460 secs.

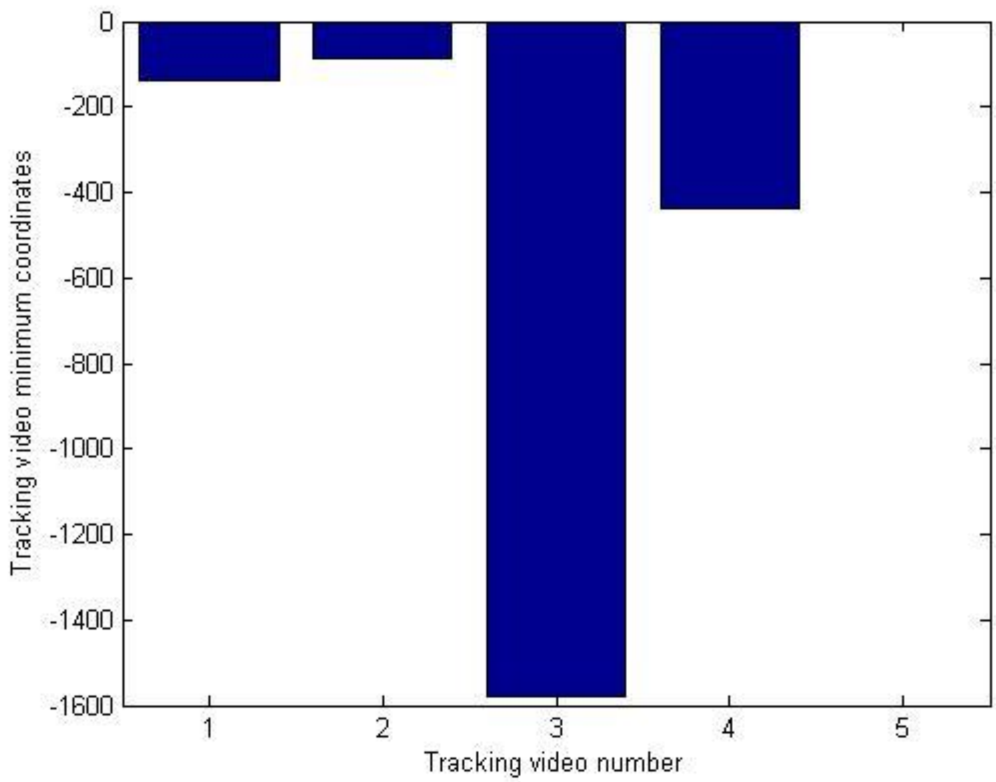


Figure 5.20: Tracking video’s coordinates

Analysis of minimum coordination of unknown objects is represented in above figure. Video 3 in figure indicates minimum coordinates level for tracking which approx 500 to 1600.

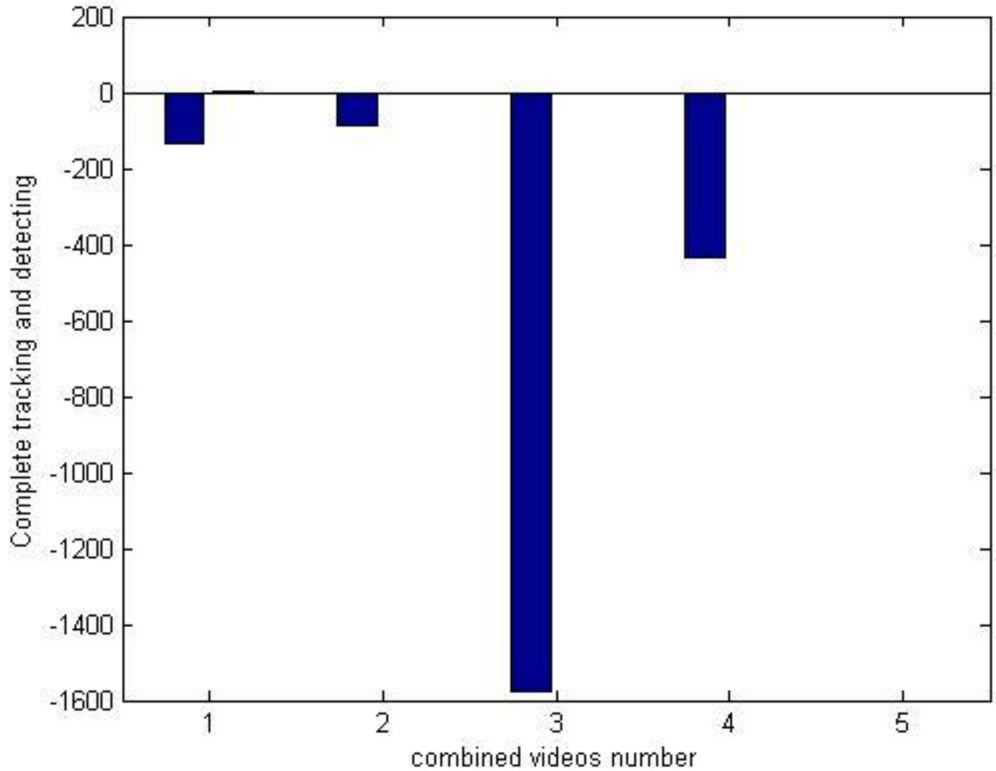


Figure 5.21: Complete tracking and detection

In the figure only video 1 show detection coordinates but other video shows detection coordinates. Detection coordinate in rest four videos indicates that object may have been lying around the corner of the video.

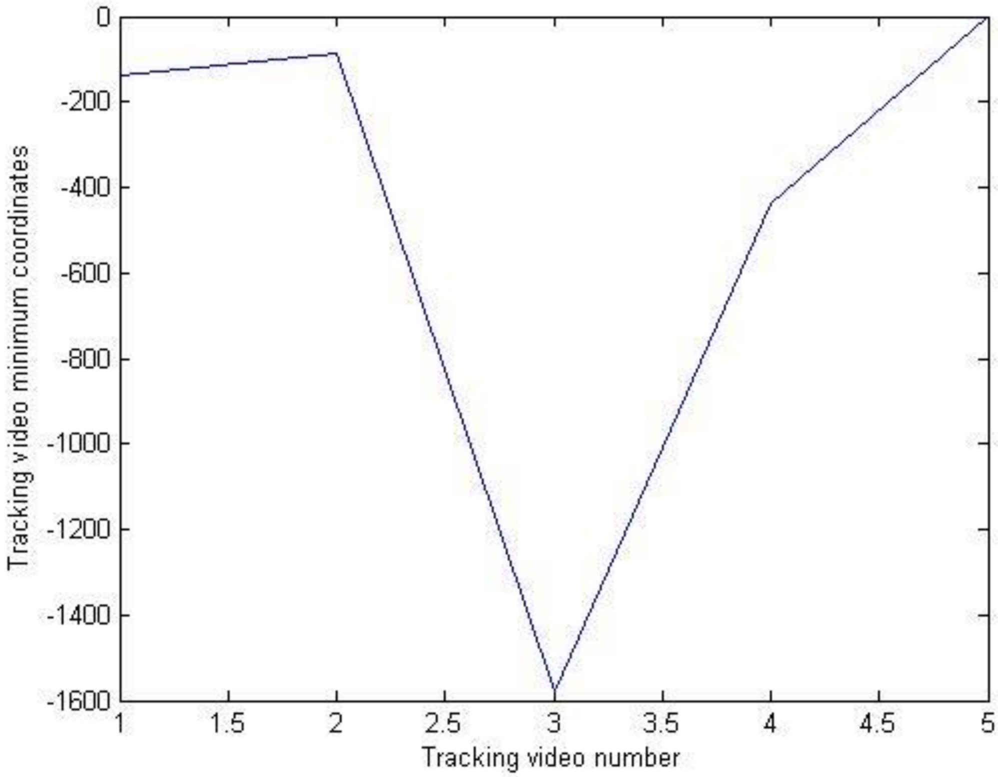


Figure 5.22: Tracking video minimum coordinates

A complete bar graph represent of detection and tracking coordinates presented.This figure represents video five is having object at the origin of this video.

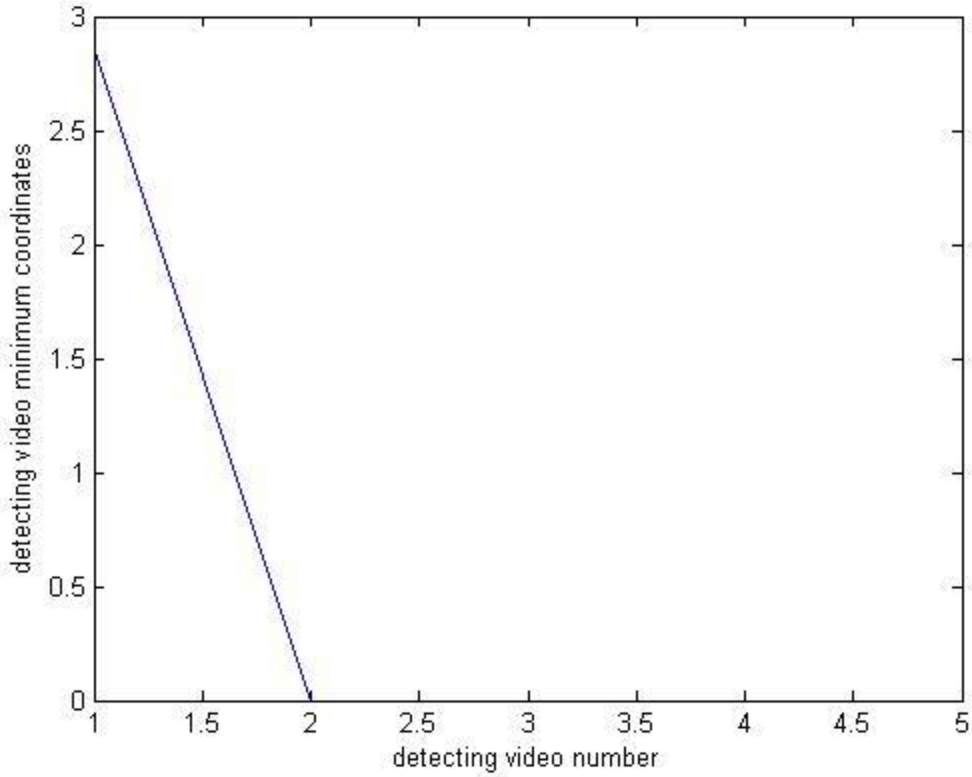


Figure 5.23: Detecting coordinates of videos

This figure is analogous as it shows a continuous plot for tracing the minimum location of coordinates among the videos.

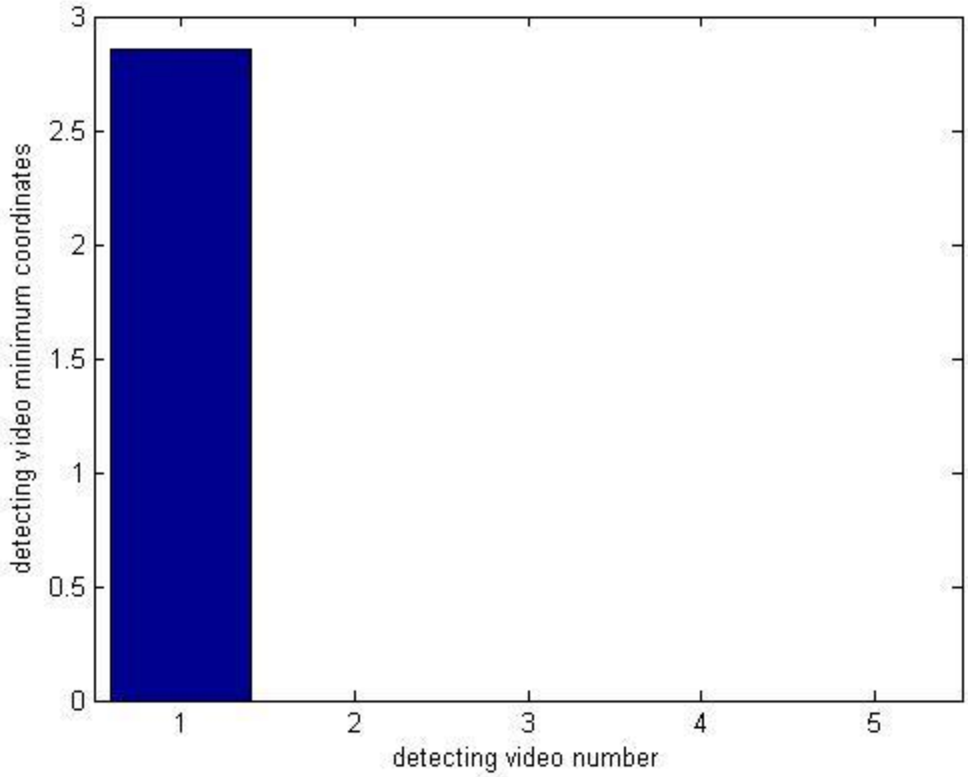


Figure 5.24: Detecting minimum coordinates

Five videos with their minimum coordinates values of detection are traced and only video 1 shows a minimum coordinate value and other videos minimum at origin.

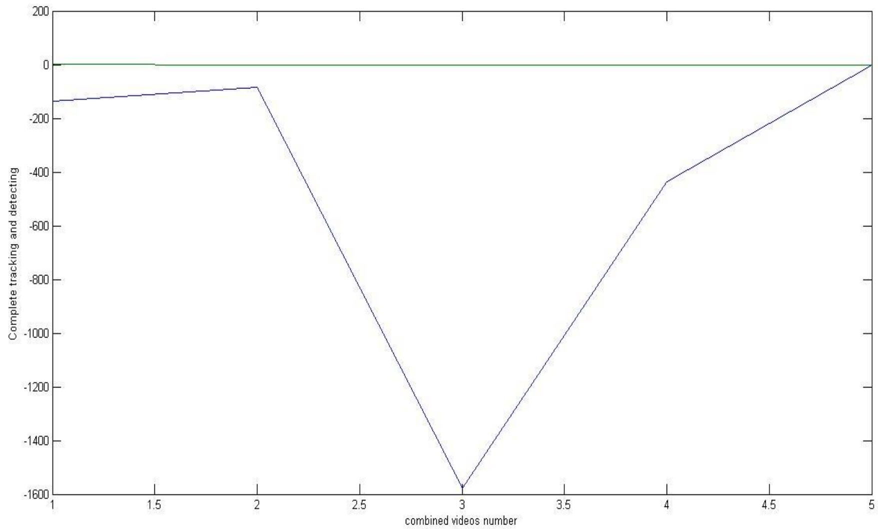


Figure 5.25: Complete tracking and detecting

Continuous plot are traced for detection and tracking in five videos and comparison of them indicates constant coordinates of detection and rapidly changing coordinates of tracking.

Chapter -6

Conclusion and Future work

6.1 Conclusion

Object detection and tracking is an important task in computer vision field. In object detection and tracking it consist of two major processes, object detection and object tracking. Object detection in video image obtained from single camera with static background that means fixing camera is achieved by background subtraction approach. In this thesis, we tried different videos with fixed camera with a single object and multiple objects to see it is able to detect objects. Motion based systems for detecting and tracking given moving object of interest can be created. Using SIFT feature extraction first feature of the object and the frame has detected to match the interested object. Since for feature extraction, SIFT algorithm has been used so tracker is invariant to representation of interested object.This thesis has examined methods to improve the performance of motion segmentation algorithms and Block matching technique for object tracking applications and examined methods for multi-modal fusion in an object tracking system. Motion segmentation is a key step in many tracking algorithms as it forms the basis of object detection. Improving segmentation results as well as being able to extract additional information such as frame difference, Gaussian of mixture model, background subtraction allows for improved object detection and thus tracking. However, a strength of kalman filter is their ability to track object in adverse situation. Integrating a kalman filter within a standard tracking system allows the kalman filter is to use progressively updated features and aids in main training identity of the tracked object, and provides tracking system with an effective means. The simulator and the simulation parameters used for the experiments.

6.2 Comparative results from base paper

According to reference paper taken for suspicious event analysis through ratio histogram is carried its detection in various video sequences by chi-hung and kuo-chin [33] is detected for various carried objects and calculated their accuracy. The result is calculated according to its original image, color of re-projection and GMM based method. The region bag detected is identified through video processing through calculating its minimum time, maximum time and processing time it needs to calculate its detection for tracking of suspicious objects. There are various videos which show suspicious objects or we can say warning shown by kalman filter by calculating its final and initial points. Table shown below shows time taken by each video for calculating its respective time for detection any object in video. The results have proved that algorithm is a robust, accurate and powerful detecting carried object detection and suspicious object analysis. Accuracy is measured by creating its GMM mixture models, segmentation, feature extraction and kalman filter. All steps required for its analysis. In our videos there is no interaction or interfering of people only there is unattended luggage in each video.

s.no	Videos	Minimum time	Maximum Time	Processing time
1.	Stairs	0	475.80	486.8982
2.	Real time Video	2.8571	859.52	91.1886
3.	VIP train	0	234.70	32.6201
4.	Unattended luggage	0	477.26	89.9048
5.	Ball	0	475.56	5.9433

Table 6.1: Interval of time taken by each video

Videos number	**New Work**		
	Abandoned objects	Inter change objects	Robbery
1.	1	0	0
2.	1	0	0
3.	0	3	0
4.	1	0	0
	Old work		
1.	74	0	0
2.	2	148	2
3.	4	8	153

Table 6.2: Accurate analysis of suspicious objects

The Table shown below shows the result of base paper calculated for the accuracy analysis of carried object detection among different event types. After carried object detection, an event classifier then can be designed for suspicious event analysis. Three event types were analyzed in this letter, i.e., “robbery,” “abandoned,” and “interchange. “List the accuracy analysis among the three event types. All the below results have proved that our method is a robust, accurate, and powerful tool for carried object detection and suspicious event analysis. Hence result is compared that is old work of base paper is compared with new work of our thesis and comparison is shown in table for videos carried out for calculating its end point of analysis through kalman filter for knowing initial and final points of any carried objects if its suspicious or not and calculating its end-points.

In the base paper they used finite state machines for suspicious event analysis if the event carried by carried bag. Generally bags transferring condition can be happen in robbery between victim and robber. The finite state machine always stays in the general state until a bag transferring state is enabled, a robber event is detected and an alarm sent. And if bag state remains at one place for long time than bag is considered as abandoned object. Hence in paper transferring and robbery both are happening for accurate analysis of bag detection. In our work we have taken videos we have taken videos to detect and track objects. Detect through X-Ray analysis, multi-staging cameras and more methods can be taken to detect objects and tracking through kalman filter of abandoned objects. Hence there is no robbery in our taken videos only interchange and abandoned objects are found.

6.3 Future Work

In the future, we can extend the work to detect the moving object with non-static background, having multiple cameras which can be used in real time surveillance applications. And various other techniques through kalman filter or Gaussian mixture model methods and algorithm to diagnose the tracking system by multiple images or objects by locating place. Various methods in improving the performance of motion segmentation algorithms and Block matching technique for object tracking applications its examined methods and multi-modal fusion in an object tracking system. Motion segmentation is a key step in many tracking algorithms as it forms the basis of object detection. Improving segmentation results as well as being able to extract additional information such as frame difference. Kalman filter is for obtaining final and initial points then detecting various correlations. More we study more we obtain. Still research is going on.

Satellite based object detection and tracking is one of the fundamental area of research. The complete analysis of such object detection and tracking can also be used in different areas for example at the mountains for tracking people doing mountaineering, on aircrafts for tracking their last known positions before air crash and in seas as well where tracking of pirate ships can be done.

Chapter-7

References

[1] AlperYilmaz, Omar Javed, and Mubarak Shah. Object tracking: A survey. Acm Computing Surveys (CSUR), 38(4):13, 2006.

[2] Gary Bishop and Greg Konolige. An introduction to the kalman filter. Proc of SIGGRAPH, Course, 8:27599–3175, 2001.

[3] J CezarSilveira Jacques, Claudio Rosito Jung, and SoraiaRauppMusse. Background subtraction and shadow detection in grayscale video sequences. In Computer Graphics and Image Processing, 2005.SIBGRAPI 2005. 18th Brazilian Symposium on, pages 189–196. IEEE, 2005.

[4] Budi Sugandi, Hyoungseop Kim, JooKooi Tan, and Seiji Ishikawa.A block matching technique for object tracking employing peripheral increment sign correlation image.In Computer and Communication Engineering, 2008.ICCCE 2008. International Conference on, pages 113–117. IEEE, 2008.

[5] Alan J Lipton, Hironobu Fujiyoshi, and Raju S Patil. Moving target classification and tracking from real-time video. In Applications of Computer Vision, 1998.WACV'98.Proceedings., Fourth IEEE Workshop on, pages 8–14. IEEE, 1998.

[6] Chris Stauffer and W Eric L Grimson. Adaptive background mixture models for real-time tracking. In Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on., volume 2. IEEE, 1999.

[7] Liu et al, Haizhou Ai, and Guang-you Xu. Moving object detection and tracking based on background subtraction. In Multispectral Image Processing and Pattern Recognition, pages 62–66. International Society for Optics and Photonics, 2001.

[8] Changick Kim and Jenq-Neng Hwang. Fast and automatic video object segmentation and tracking for content-based applications. Circuits and Systems for Video Technology, IEEE Transactions on, 12(2):122–129, 2002.

[9] Shahbe Mat Desa and Qussay A Salih. Image subtraction for real time moving object extraction.In Computer Graphics, Imaging and Visualization, 2004.CGIV 2004.Proceedings. International Conference on, pages 41–45. IEEE, 2004.

[10] Budi Sugandi, Hyoungseop Kim, JooKooi Tan, and Seiji Ishikawa.Tracking of moving objects by using a low resolution image.In Innovative Computing, Information and Control, 2007.ICICIC'07. Second International Conference on, pages 408–408. IEEE, 2007.

[11] YUTAKA Satoh, S Kaneko, and SATORU Igarashi. Robust object detection and segmentation by peripheral increment sign correlation image. Trans. of the IEICE, 84(12):2585–2594, 2001.

[12] Mahbub Murshed1/2, MdHasanul Kabir1/2, and Oksam Chae1/2. Moving object tracking-an edge segment based approach. 2011.

[13] Weiming Hu, Tieniu Tan, Liang Wang, and Steve Maybank.A survey on visual surveillance of object motion and behaviors. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 34(3):334–352, 2004.

[14] Zhan Chaohui, DuanXiaohui, XuShuoyu, Song Zheng, and Luo Min. An improved moving object detection algorithm based on frame difference and edge detection. In Image and Graphics, 2007.ICIG 2007. Fourth International Conference on, pages 519–523. IEEE, 2007.

[15] Ismail Haritaoglu, David Harwood, and Larry S. Davis. W¡ sup¿ 4¡/sup¿: real-time surveillance of people and their activities. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 22(8):809–830, 2000.

[16] Deep J Shah, Deborah Estrin, and AfrouzAzari. Motion based bird sensing using frame differencing and gaussian mixture. Undergraduate Research Journal, page 47, 2008.

[17] Budi Sugandi, Hyoungseop Kim, JooKooi Tan, and Seiji Ishikawa.Tracking of moving objects by using a low resolution image.In Innovative Computing, Information and Control, 2007.ICICIC'07. Second International Conference on, pages 408–408. IEEE, 2007.

[18] IntanKartika and ShahrizatShaik Mohamed. Frame differencing with post-processing techniques for moving object detection in outdoor environment. In Signal Processing and its Applications (CSPA), 2011 IEEE 7th International Colloquium on, pages 172–176.IEEE, 2011.

[19] Robert T Collins, Alan Lipton, Takeo Kanade, Hironobu Fujiyoshi, Rosals and scalaroff, YanghaiTsin, David Tolliver, Nobuyoshi Enomoto, Osamu Hasegawa, Peter Burt, et al. A system for video surveillance and monitoring, volume 102. Carnegie Mellon University, the Robotics Institute Pittsburg, 2000.

[20] Chris Stauffer and W Eric L Grimson. Adaptive background mixture models for real-time tracking. In Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on., volume 2. IEEE, 1999.

[21] PakornKaewTraKulPong and Richard Bowden.An improved adaptive background mixture model for real-time tracking with shadow detection. In Video-Based Surveillance Systems, pages 135–144. Springer, 2002.

[22] Cl´audioRosito Jung. Efficient background subtraction and shadow removal for monochromatic video sequences. Multimedia, IEEE Transactions on, 11(3):571–577, 2009.

[23] Yung-Gi Wu Cheng and Chung-Ying Tsai Chen. The improvement of the background subtraction and shadow detection in grayscale video sequences.In Machine Vision and Image Processing Conference, 2007.IMVIP 2007. International, pages 206–206. IEEE, 2007.

[24] Muhammad Shoaib, Ralf Dragon, and JornOstermann.Shadow detection for moving humans using gradient-based background subtraction.In Acoustics, Speech and Signal Processing, 2009.ICASSP 2009. IEEE International Conference on, pages 773–776. IEEE, 2009.

[25] Jianhua Ye, Tao Gao, and Jun Zhang. Moving object detection with background subtraction and shadow removal.In Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on, pages 1859–1863.IEEE, 2012.

[26] T Thongkamwitoon, S Aramvith, and TH Chalidabhongse. An adaptive real-time background subtraction and moving shadows detection.In Multimedia and Expo, 2004.ICME'04. 2004 IEEE International Conference on, volume 2, pages 1459–1462. IEEE, 2004.

[27] PakornKaewTraKulPong and Richard Bowden.An improved adaptive background mixture model for real-time tracking with shadow detection. In Video-Based Surveillance Systems, pages 135–144. Springer, 2002.

[28] John Brandon Laflen, Christopher R Greco, Glen W Brooksby, and Eamon B Barrett.Objective performance evaluation of a moving object super-resolution system. In Applied Imagery Pattern Recognition Workshop (AIPRW), 2009 IEEE, pages 1–8. IEEE, 2009.

[29] Lloyd L Coulter, Douglas A Stow, Yu Hsin Tsai, Christopher M Chavis, Richard W McCreight, Christopher D Lippitt, and Grant W Fraley. A new paradigm for persistent wide area surveillance. In Homeland Security (HST), 2012 IEEE Conference on Technologies for, pages 51–60. IEEE, 2012.

[30] ShaliniAgarwal and Shaili Mishra.A study of multiple human tracking for visual surveillance.International Journal, 5, 1963.

[31] FatihPorikli. Achieving real-time object detection and tracking under extreme conditions. Journal of Real-Time Image Processing, 1(1):33–40, 2006.

[32] Huchuan Lu Cozy, Ruijuan Zhang, and Yen-Wei Chen. Head detection and tracking by mean-shift and kalman filter. In Innovative Computing Information and Control, 2008.ICICIC'08. 3rd International Conference on, pages 357–357. IEEE, 2008.

[33] Chi-Hung Chuang, Jun-Wei Hsieh, Luo-Wei Tsai, Sin-Yu Chen, and Kuo-Chin Fan.Carried Object Detection Using Ratio Histogram and its Application to Suspicious Event Analysis IEEE transactions on circuit and systems for video technology vol. 19, no. 6, June 2009.

[34] Bastian Leibe, Konrad Schindler, NicoCornelis, and Luc Van Gool. Coupled object detection and tracking from static cameras and moving vehicles. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 30(10):1683–1698, 2008.

[35] Bastian Leibe, Konrad Schindler, NicoCornelis, and Luc Van Gool. Coupled object detection and tracking from static cameras and moving vehicles. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 30(10):1683–1698, 2008