SUSPICIOUS HUMAN ACTIVITY DETECTION FROM SURVEILLANCE VIDEOS

FATHIA SALEM

OCTOBER 2017

SUSPICIOUS HUMAN ACTIVITY DETECTION FROM SURVEILLANCE VIDEOS

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF NATRUAL AND APPLIED

SCIENCE OF

ÇANKAYA UNIVERSITY

BY

FATHIA SALEM

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS THE

DEGREE OF

MASTER OF SCIENCES

IN

COMPUTER ENGINEERING

OCTOBER 2017

Title of the Thesis: **SUSPICIOUS HUMAN ACTIVITY DETECTION FROM SURVEILLANCE VIDEOS.**

Submitted by **FATHIA SALEM**

Approval of the Graduate School of Natural and Applied Sciences, Çankaya University.

Prof. Dr. Can ÇOĞUN
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Erdoğan DOĞDU
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Reza HASSANPOUR
Supervisor

Examination Date: 16.10.2017

Examining Committee Members

Prof. Dr. Mehmet TOLUN (Aksaray Univ.)

Assoc. Prof. Dr. Reza HASSANPOUR (Çankaya Univ.)

Assist. Prof. Dr. Abdul Kadir GÖRÜR (Çankaya Univ.)

STATEMENT OF NON-PLAGIARISM PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Fathia Salem

Signature :

Date :16.10.2017

ABSTRACT

SUSPICIOUS HUMAN ACTIVITY DETECTION FROM SURVEILLANCE VIDEOS

SALEM, Fathia

M.S., Computer Engineering Department

Supervisor: Assoc. Prof. Dr. Reza HASSANPOUR

October 2017, 104 pages

Video surveillance has been used from a long time to provide security in many sensitive places, so with this great progress in various aspects of life the traditional surveillance operations are facing many problems because of the large amounts of information that must be handled manually in a limited time also the possibility of information loss which can contain important things such as suspicious behaviours. So recently, a large amount of research has been conducted on video surveillance.

In this thesis, we will present a system to support the smart surveillance for detecting abnormal behaviours that represent security risk. The proposed algorithms are intended to detect two cases of human activities namely, walking and running. We impose no restriction on the number of people in the scene, and the direction of the motions. However, we restrict the videos to indoor colour videos, where the video are captured by one stationary camera. The moving objects which correspond to people in the scene are detected by background subtraction algorithm. We consider the displacement rate of the centroids of the segmented foreground areas and the rate of change in the size of the segmented areas as the two main features for activity

classification. The proposed algorithms determine the activity type with a high accuracy rate.

Keywords: Video Surveillance, Suspicious Human Behaviour, Security.

ÖZ

GÖZETİM VİDEOLARI İLE ŞÜPHELİ İNSAN FAALİYETLERİNİN BELİRLENMESİ

SALEM, Fathia

Yüksek Lisans, Bilgisayar Mühendisliği Anabilim Dalı

Tez Yöneticisi: Doç. Dr. Reza HASSNPOUR

Ekim 2017, 104 sayfa

Video gözetimi birçok hassas yerde güvenliği sağlamak için uzun zamandan beri kullanılmaktadır ve hayatın çeşitli yönlerinde meydana gelen bu büyük ilerleme nedeniyle, geleneksel gözetim faaliyetleri sınırlı bir alanda ele alınması ve kullanılması gereken büyük miktarlardaki bilgiler yüzünden birçok problemle karşı karşıyadır ve şüpheli davranışlar gibi önemli hususları içerebilen bilgilerin kaybedilmesi olasılığı da dikkate alınmalıdır. Son zamanlarda, video ile gözetim konusunda çok sayıda araştırma yapılmıştır.

Bu tezde güvenlik riskini temsil eden anormal davranışların saptanması için akıllı gözetimi destekleyecek bir sistem sunacağız. Önerilen algoritmalar, iki insan faaliyetini, yani yürüme ve koşmayı tespit etmeyi amaçlamaktadır. Olay yerindeki kişilerin sayısına ve hareketlerin yönüne hiçbir kısıtlama getirmeyeceğiz. Bununla birlikte, görüntünün sabit bir fotoğraf makinesiyle yakalandığı kapalı renk videoları kısıtlıyoruz. Olay yerindeki insanlarla uyumlu olan hareketli nesneler geri plan çıkarım algoritması tarafından algılanır. Biz faaliyet sınıflandırması açısından iki önemli özellik olarak, bölümlere ayrılmış ön alanların ağırlık merkezinin yer değiştirme oranını ve bölümlere ayrılmış alanların boyutundaki değişim oranını dikkate alıyoruz.

Önerilen algoritmalar, etkinlik türünü yüksek doğruluk oranı ile belirler.

Anahtar Kelimeler: Video Gözetimi, Şüpheli İnsan Davranışı, Güvenlik.

ACKNOWLEDGMENT

"Never give up on a dream just because of the time it will take to accomplish it. The time will pass anyway." — Earl Nightingale

"I am really thankful to my GOD for showing the right path to me"

To spirit of my father and my mother, I present my success to them and how much I wish they would be with me at this moment, God's mercy on them.

I am deeply indebted to everyone who contributed to the completion of this dissertation even though a simple words, Foremost, I want to express my sincere gratitude for my Supervisor Assist. Prof. Dr. Reza Hassnpour for all his advices and guidance, which were very deep and inspirational to my thesis, which without it I would never have reached this point.

My special thanks for my dear husband "Rafi" to support me against all the difficulties with me, to my children little prince and princess "Hussam", "Pinar" and my new angel "Al-zobair", who gave me a moral support to continue in weakness times.

I give my great thanks to my family brothers and sisters (Hamida, Hammad, Moftah, Sameer, Mlla, Saadona) for support me in all my different stages life.

Finally, I thank all my friends who encouraged me as much as possible.

TABLE OF COTENTS

LIST OF FIGURES

FIGURES

LIST OF TABLES

TABELS

LIST OF ABBREVATIONS

HOG	Histogram of Oriented gradient.
SVM	Support vector machine.
CAMSHIFT	Continuously adaptive mean shift.
DE	Displacement error.
ROI	Region of interest.
GMM	Gaussian Mixture Model.
HOOF	Histogram of oriented optical flow.
MFPM	Mean Feature Point Matching algorithm.
SURF	Speeded-Up Robust Features method.
MSER	Maximally Stable Extremal Regions.
DOG	Difference of Gaussian.
ADI	Absolute value of the differential in image.
ART	Angular radial transform.
EBD	Entropy Based Discretization.
DOF	Difference Of Frames.
DBN	Dynamic Bayesian Networks.
SIFT	Scale Invariant Feature Transform trajectories descriptor.
MEI	Motion Energy Images.
MHI	Motion History Images.

STIP	Spatio-Temporal Interest Point.
HOF	Histogram Optical Flow.
BOW	Bag of Word.
UTD-MHAD	Multimodal Human Action Dataset.
ATS	Adaptive Temporal Sampling.
DMMs	Depth Motion Maps.
DSM	Depth Static model.
PCA	Principal Component Analysis.
MSR	Microsoft Research.

CHAPTER 1

INTRODUCTION

In recent times, the world has become dependent on multiple types of applications, which serve us in all areas of life. One of most important of these applications are the video surveillance systems [1, 2, 3, 20, and 21]. Therefore, these systems play an important and effective role in providing security and guarantee it to individuals. Therefore, this project is important in this area to reduce the work of the riots and the deployment of security as much as possible

We can present video surveillance system as analysis of the consecutive video frames to search for suspicious activities and detect them. Traditional surveillance systems can be done manually "semi-autonomous" this means by security individuals, but this way is complicated, expensive and the accident rate is high, so the best way is fully autonomous system or a video surveillance system.

The process of detecting moving objects in video is one of the basic steps in the task of video analysis figure 1, where the moving objects are detected from the stationary background by using certain techniques. These techniques are: “Background Subtraction Method [4, 5, 6, and 7]” ,“Optical Flow Method [8,9,23,24]” and ”Frame-To-Frame Difference Method [10,11,25,26]”. In this research we will use “Background Subtraction Method” to detect the moving objects from the stationary background. There are many challenges in this project these challenges occur due to changes in environmental conditions like lighting, reflections and shadows, so the segmentation of objects is a difficult issue and need to be dealt with well by using a robust surveillance system, we reduced it by used the morphological operations to reduce the noise. Also there is an important step which is object classification where this step is done by using two main approaches, "shape-based classification" and "motion-based classification", where spatial information is used in "Shape-based methods", as for "Motion-based methods" they use temporal information to classify objects.

After object classification phase the next step in this system (video analysis sequence) is detect the activity and determine if it is suspicious or non - suspicious and this step applies to the following:

1. For the objects that are moving parallel with the camera, we have to find value of centroid to know if this object (human) is moving from right to left or from left to right, we do this by calculating the absolute value of the x coordinate of centroid for consecutive frames, and compare the displacement value in x with the threshold value to know if this activity is normal (walking) or suspicious (running).
2. For the objects that are moving (To/ away) from the camera we compute the change area for object size between consecutive frames, if the change is small that means the activity is normal (non-suspicious/walking), vice versa if the change is big that mean the activity is suspicious (running).

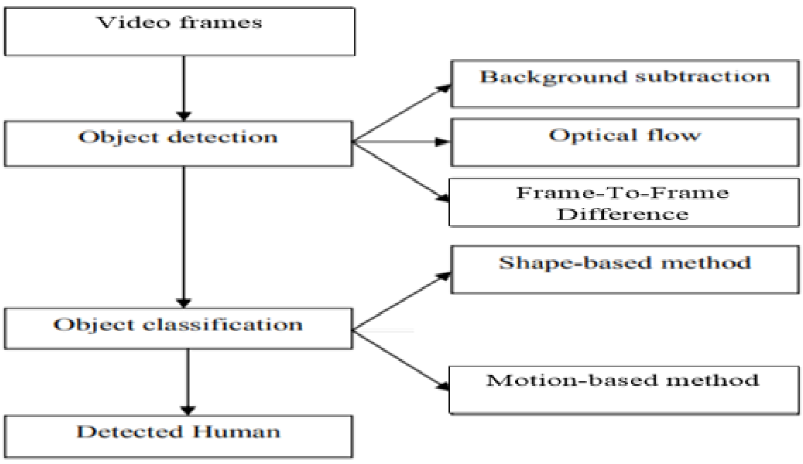


Figure 1: Object detection.

1.1 Motivation

Video surveillance systems are playing an important role in providing security and ensuring it in all places, whether private or public this is the main aim for it, where this detects the suspicious activities such as running ,fighting and reduce their occurrence, thus providing security and preventing any incidents that may disturb security. Therefore, minimizing terrorist operations that could occur.

1.2 Scope of Thesis:

This system focuses on detecting the suspicious activities by using Background subtraction method for object detection and uses the morphological operations to reduce the noise that results from natural changes like lighting and shadows, where the input video will be indoors and it will be captured by one fixed camera in MOV format with 23 frames per second, the size of the object will range from 1-2 meters (taller than 1 meter , shorter than 2 meters), the distance between the object and the camera will range from 4-5 meters for near cases and 9-10 meters for far cases. The video will contain an object or two as a maximum; furthermore our data set has a faint lighting and does not have any reflections.

1.3 Organization of the Thesis:

The organization of this thesis is as following:

Chapter 1 presents the introduction where the problem is presented, how we solved it, and the motivation (why we use this system?). Finally, scope of this thesis.

Chapter 2 presents the literature survey we summarize the work done by other researchers in this field.

Chapter 3 presents the proposed method where we described and gave all details of our method.

Chapter 4 presents the experimental results where we discussed the results that we got.

Chapter 5 presents the conclusion and the future work where we provide a short summary about our work, also we presented how our methods can be extended and which parts can be improved.

CHAPTER 2

LITERATURE SURVEY

2.1 Introduction

Automatic determination of human activity type has attracted a vast number of researchers in the recent years. This trend has been intensified by the recent security threats and the need for real-time analysis of the surveillance videos. In this thesis, we will focus on the abnormal detection in surveillance videos and all things related to this subject, especially the abnormal detection that related to running indoor, because this is the most dangerous activity that most likely will lead to a security disorder. Many different approaches address the problem of suspicious human activities [4, 5, and 8]. These approaches can be categorized as:

1. Object detection.
2. Object classification.
3. Object tracking.

In the following sub-sections, we describe main methods from each group.

2.2 Object Detection:

The majority of systems that work based on visual surveillance systems begins with Motion Detection, there are many methods or techniques that are used for Motion Detection these methods depend on trying to locate the regions of pixels which reflect the moving objects in the scene, These methods are summarized in *"Frame-To-Frame Difference Method"* [10, 11, 25, 26], "*Background Subtraction Method"* [4, 5, 6, 7], *"Optical Flow Method"* [8, 9, 23, 24]. Where we can say the desired result of the motion detection is segmenting the corresponding areas for the objects that move from the rest of the image, concept of motion and object detection is always based on background models.

2.2.1 Background Subtraction :

Background subtraction [4, 5, 6, 7] is one of the methods that is characterized by a smoothness apply so it enjoys a wide popularity for applications that have static backgrounds, detecting the moving areas in the image by taking the difference between the current image and background image reference by using a pixel-by-pixel way, this method is very sensitive to light and its changes.

- The authors in this paper [4] worked on the recognition of human activity and behavior and in the final step get important information to the observers who are responsible for the internal control system. Also, we shall notice that the worker group on this paper interested to using "*background subtraction algorithm"* for the purpose of multi-object detection ,and they used "*HOG feature"* and "*SVM classifier"* to recognize the human in the image, on the other hand they used "*Viola–Jones algorithm"* to capture the face of the person they want to detect, finally they used "*Mean Shift Technique"* to detect the behaviors and this happens by tracking the person based on their individual appearance, this algorithm is a method for non-parametric clustering that means the foreknowledge of the number of clusters is not required, "*Mean-ratio"* and "*Log Ratio operators"* are used to change detection.

The steps of System Model it was as follows: a) Video Input, b) Background Image Acquisition: In this section they based on capture the frames as the reference for images to use these references for any additional processing, moreover the background is set only once, where the camera is programmed to capture direct video specifically for observation region, c)Image pre-processing: Here are the images that were obtained are pre-processing in order to improve its frames, as we know that the video frames contain a lot of noise due to several reasons such as shadows and lighting where must remove this noise to be getting good results and this is done at this stage pre-processing, After several days in the digital domain is performed pre-processing video process which is done after the digital video capture, d) Change Detection: They used Mean-ratio operators for change detection, as we know in the past step we stored the reference of frames which are took and have been processed, In this step the video that was taken in order to be detected any change with the reference frame as a reference condition. Where it is to use the tools available within the MATLAB environment to handle images that have been obtained so that the control of the camera, which is connected directly with the computer where operated those functions to get the shots required of the camera, after obtaining all required snapshots which is actually is the input video then the separation foreground images from the background images is achieved at this stage.

The authors also discussed foreground extraction: the backgrounds that were obtained in previous steps that was considered as a reference image so as to be ready for image segmentation and here are rebuilding the foreground object, by removing the background elements. Where it is converting images from the colored images to grayscale because work with grayscale images 2-D is much easier than work with color images (RGB) 3-D especially to compare a pixel to pixel for identifying the foreground images. As well as the noise has been removed from the images by using morphological operations so that we get a binary dilated image which is used structuring elements which return the dilated image and this enables them to collect a large amount of information in a short time. Also, this leads to the reduction of time that spent in processing and thus use the time for more image processing operations.

On the other hand they wrote about Change Detection, the sequential steps algorithm of this algorithm are as following: 1) When starting in processing a new image, we note the probability of pixel value in the new image will be equal to pixel value in background image, 2) If the value of pixel is greater than the value of the allowable range (threshold), then the pixel is stored and put in as a mark to indicate that it is a part of the foreground image. This is based on if this pixel belongs to the background image and that is through the allowable range is also expected that this pixel will be repeated for a long period of time, 3) If it was the pixel which was observed within the threshold value, this means that is part of the background image and is replaced it with a value of zero. As well as if there was a major change in pixel value, it is kept and considered as part of the change in the foreground.

As for that abandoned object detection, the authors here discussed this issue based on the timer idea, which starts to work in case of detecting the presence of the object in the image and grows significantly until it finds the static object. Compared the incremented timer against threshold timer value which has been defined previously, if the incremented timer value equals or exceeds the threshold value until just once it means this object will be ignored. Moreover, the next step is to raise an alarm about an abandoned object as well as identifying the object by using a rectangle on the screen to focus on the area which has been monitored. Also they indicated to activity analysis this activities include running, walking, jumping and bending: for *running* this based on the speed when the change in value of X-axis across the subsequent frames exceed the range of threshold value, *walking* But if the value of speed was less than features of threshold it is considered as a normal walk, as for *Jumping* this is based on change of Y-axis value Furthermore, the speed of this up and down motion should be exceed the range of threshold value, *bending* can be detect about this activity when the aspect ratio of the detected object decreases across the frames exceed the range of threshold value.

Where the distance has been calculated by using formula of centroid, the value of variables are the positions of pixel for the person from first stage to final stage:

$$\text{Distance} = \sqrt{(x2 - x1)^2 + (y2 - y1)^2} \qquad (2.1)$$

Where,

X1: previous pixel position.

X2: pixel position in width.

Y1: previous pixel position.

Y2: pixel position in height.

They can be determined the speed of moving the object by using the distance of travelled by the centroid to the frame rate of the video.

Velocity= Distance travelled/Frame rate.

If we want to discuss the advantages of the proposed algorithm which is background subtraction algorithm with Mean Shift technique and Thresholding, by using the morphological operations they remove the noise from images so they only deal with important information in the image therefore reducing the processing time, we can note from the final results it is effective with regard to detection of the object abandoned and the abnormal behaviors in safe areas, where the result of walking was Motion_X=10, Velocity=10.1783, Motion_Y=4, Aspect ratio=2.7213, So it was classified as walking because the value change of X-axis is too small, while the result of running (abnormal behavior) was Motion_x=169, Velocity=131.4364, Motion_Y=68, Aspect ratio=3.8769 classified as running because the1 value of X-axis across the subsequent frames exceed the range of threshold value, for bending was

Motion_x=3, Velocity=13.3417, Motion_Y=13, Aspect ratio=0.03333 classified as bending because aspect ratio across the frames exceed the range of threshold value, and jumping was Motion_x=12,Velocity=42.72, Motion_Y=41, Aspect ratio=3.0946 classified as jumping because aspect ratio decreases across the frames exceed the range of threshold value.

For disadvantages I think that there is a congestion of methods which have been used and thus can be difficult to control the results and to get a clear result after the implementation of all these methods at the same time, also this system is very sensitive for lighting also it needs to deal with removing noise from all frames to get clear backgrounds and this will be difficult.

- In this paper [5], the proposed system offers an "*adaptive subtraction for background"* which is also consistent with differentiating for foreground extraction this done by extracted map of foreground from background model during the processing. Afterwards, it proposes an elimination phase to deal with shadow, light and labelling problem. Beside this, the model uses a colour correlogram for tracking. As for the terminological part of the proposal, we can firstly mention about the term morphological operation. Simply put, morphological image processing means operations which has been carried out such as erosion and dilation in order to remove noise and repair deadlock from the extracted foreground where used an array with arbitrary size which is a components of 0 and 1 it has name a template. Secondly, colour histogram means the tonal dispersion in the video. Lastly, colour correlogram represents a term calculating the tonal dispersion in the video as well but here it is has another attribute that is spatial information, so the process of description is more clarity and accuracy.

The method of this model is carried out by several phases. First, "a hybrid background subtraction algorithm" does the foreground detection. Afterwards shadows are removed and it is followed by morphological operations. Then the noise elimination phase and labelling are performed respectively. Lastly, the whole process is ended by the tracking phase. In the first phase, known as the detection of foreground, the model uses *"a hybrid background subtraction algorithm"*. This algorithm works by subtract pixel *x* from the foreground and pixel *x* from existing frame if difference density is more than existing threshold this mean the model of background it will update for every pixel followed for non-foreground in the existing frame, thus it can handle with changing of gradual lighting which is usually use for define the updating rate. Also there is other method used for same aim which is temporal differencing as follows the detected of pixel *x* is done basis on this pixel belongs to foreground if density of difference between the current frame and previous frame is higher than the certain threshold, thus here the background model is not really important, this model is effective to changes in the dynamic scenes, where it is the combination of the previous methods to give a better results for segmentation. They also used a new better technique that is exempted parameter of standardization process.

The model aims to calculating brightness and chromatic disorder, which are calculated on the assumption that I(x)=[$I_R(x)$,$I_G(x)$,$I_B(x)$] based on RGB which is a mixture of pixel x that from existing image also B(x)=[$B_R(x)$,$B_G(x)$,$B_B(x)$], those pixels corresponding in background of image, where $^{\delta}$ B(x) is brightness distortion, that clarifies result of subtraction between B(x) and I′(x), I′(x) is projection of I(x) on B(x). This algorithm allows us to repair the pixel differences between the frames in the video, and provides us information to solve illumination change to differentiate foreground from background. One of the most important interventions to recognize the objects is to cope with noise and light problems. For this, morphological applications are carried out by the model to remove noise and other occlusions from foreground objects where dilations then apply erosions after that apply dilations process again on elements of the various structuring which got it in training phase, as a result of these operations the holes that in image get corrosion also for the discontinuities in image get filled. The last thing in this phase is to use the results that have been gotten in phase of shadow detection that to detect about all wrong pixels in foreground that is happened as a result of highlighting, also the writers are used the graph algorithm to express the different objects during the all operations by using the different shapes like a rectangle or masses centres...etc.

As for the tracking phase, the model firstly uses centre of mass technique to remove the mismatch between the previous and present objects. The *colour correlogram* is used to multiple matched candidates where used to obtain spatial information and brings much more qualified information and results about objects thanks to the distance calculation technique of pixels. In case of intense occlusion, the model applies a calculation period to detect and repair occlusions that emerge as problems for the continuity of the video. Where a centre of mass is a rate of x'th, y'th for put a rectangle about the object. A correlogram indicate to possibility of presence pixel for object with specific colour, let's say that $\boldsymbol{c_i}$ is a different pixel located at a specific distance, k is a specific colour, colour correlogram is indicate to spatial information which indicate to density of the colour ,reverse the colour of histogram. As a result, we note the field of colour correlogram gives us good results. The measured the distance between the two images is done by ***I*** and ***I'***. Background model, a correlogram colour and the histogram are need to the process of updating. Therefore, the models of objects can handle with the difference in distortions or lighting...etc.

The researchers also discussed Occlusion Handling, were said, it is important to detect on events of occlusion, where it can be said that the occlusion occurred when the number of objects increased or there are any frame containing a number of surplus objects which overlap more than the previous object there are named occluded objects properly, every pixel is belong to this collection it will be computed through calculate the histogram for ratio back-projection also for correlogram correction factor.

For Experimental Results and Analysis phase, the same sample sequences they used standard "*Continuously adaptive mean shift* (CAMSHIFT) "to keep the processing and tracking the results. They compared between different data such as Ground Truth Data (x, y co-ordinates), CAMSHIFT results," Hosain and Saha" results, and current system results, they used 25 frames, where DE is" Displacement error", ROI is "Region of interest", the results show us the average DE of CAMSHIFT is 30, and average DE which is describe through "Banerjee and Sengupta" is 2.62, the average DE of current system is 1.74. So these results show us the current system is better than others.

As a result of comparison with two other works, the model's works remain more accurate in terms of the distance between the ground truth data and the tracking centre of location of interest. Additionally, even though one of the two works has used a colour correlogram, this model's systems work more effectively compared to that work, since a hybrid algorithm had already detected the foreground and eliminated shadow before. The results also show that this model is much more successful in detecting occlusion problems and solving them after splits in various environments such as indoor and semi-outdoor milieus. All in all, the results show that the system of this model offers much more accurate results coping with changing environments, lightning problems, occlusions, rotations, splits and mismatch between previous and present objects, thanks to its new algorithm approach and colour correlogram technique which reduces the possible mismatch and the possibilities of other occlusions in the video.

To sum, this model offers a solid and faster approach for video surveillance issue. Therefore, the model applies two various algorithms which are the adaptive technique and robust method to detect foreground and shadow. In the tracking phase, the model uses a technique with colour correlogram and histogram. The model aims at reducing distance measures to get the correct match between objects by using more effective correlogram. However, the model cannot fairly cope with umbra shadow and it assumes background totally static while dealing with penumbra shadow in a great way. Yet, for this problem, this model also offers that changing background can be solved by means of analysis of temporal variations. Therefore, the model aims at widening future works in order to cope with changing background problem and develop much better performance for tracking phase. But there are some disadvantages in this paper which are in Feature extraction phase they are not covered sufficiently it has been necessary to identify those features such as height or width or ... etc. This algorithm is not effective with complex scenes also when the background has similar conditions with the objects so it does not give us the desired results Also about Kalman filter phase this phase is important and the writers did not discussed it sufficiently.

- The researchers in the proposed algorithm of [6], they have discussed the subject of the human behavior detection based on "*Background Subtraction Algorithm*", also they apply threshold segmentation technique, differencing technique, morphological operations technique, and object tracking technique, all these techniques apply in real time.

Where they discussed the moving object detection as an additional process this is occur after classified it in the real time development application, the big challenge here is detection of objects within a period of time without any pressures by using available hardware and high efficiency where background subtraction process occur by input the stream video in real time to process this stream, generation this video and discrimination each frame and find the threshold for tracking objects, when capturing a new image they compute the differential between background and the image for moving object detection but this process is a complex so they used another approaches

to perform this process effectively, Note that detection the moving objects are used heavily and this operations for the purpose of developing real-time applications which are used in the process generation of surveillance with high efficiency, as we know in each research there are many challenges the moving object detection is one of difficult challenges especially in applications of real time visualization system ,With regard to detect object categorization worthily there is a main technique which is "*edge localization"*, it is used "a *gradient operator"* to create a map of images gradient which is consisting of inputs and background images where is calculated the map of gradient difference from map of gradient images also it is detected the moving object by using "*masking of suitable directional and threshold"*, when we focus on the processing of movies in real time applications we can conclude the human can move in any direction with semantic or normal behavior also he can do some unnecessary activities which is provide in video processing this categorization, might happen defect between the natural movement of the human body with the factual information event generation also the activity anomaly can occur in video surveillance with things other unwanted there realistic generation in cases of processing.
For motivation they said there are many current techniques that are used for video surveillance which are efficient in categorizing the objects, but they also pointed out that those techniques did not detect all things that are unnecessary in the human body, so it was necessary to reach to technique for detection of suspicious activities which relevant to the development applications in real time and this is considered one of the important *challenges* in this paper.

On other hand they proposed frame subtraction method to detect categorization of the objects, is being used it has been proposed by an algorithm by "Widyawan Muhammad" which is adaptive motion detector, also it has been used by the frame difference method which uses a special technique to select which is a reference image it will use for the motion detection, it is known that the technique is like template matching, there are two methods for this technique which are semantic method and feature method they used to develop the actions that not necessary in real time when the template matching is done successfully that means the event progression is restored with "*data event generation*" ,they also used *Optical flow* method that for estimate the motion of objects during the string of frames, that method depends on the values of pixel when the points are located on same object, the values of the object pixel with same location it will have constant brightness all the time, the writers also considered the optical flow have two Ingredients first one is normal flow as for the second one is parallel flow, “Deval Jansari” and “Sankar Parner” suggested an optical flow method for progression the data that in real time, in this method is applied some of procedures to obtain the region of changes, these procedures are subtracted $I(x,y,t)$, $I(x, y,t + \Delta t)$ and applied thresholding on the difference frame ,but there are some problems in this method like sensitive to noise, the vast amount of calculations, low level performance in the resistance the noise during the real-time. In this paper also they discussed background propagation method which is based on computation of all frames, where it has been calculated density of all existing pixels in each frame so as to reach to the background frame, here is being taken the difference between the existing image and the reference image that is very sensitive so they proposed method for object moving detection that is “*background subtraction method”*.

For experimental results they capture two frames that are sequential (N,N+1), time between the frames is limited, converted all frames from color scale to gray scale, after that is subtracted N frame from N+1 frame that to get the difference image also they used “*Sobel filtering”* and applied it on the difference image for detecting edges of the image and removing the noise, moreover, “*median filtering”* been applied to reduce the probability noise, here in current technique it marks the location of object frames (N,N+1) ,they calculated the moving object if it is fast this means distance it will be large, (slowly move, small distance), they show us some results like for each spatiotemporal template it takes less than 0.06 seconds for each frame 2.8 GHz ,when they used 432 templates they got time reached to 25 seconds for processing each frame. In this paper there are many disadvantages which are there are some false positive cases and error rate is exist in the results, the results it is not discussed enough to figure out what the system was working efficiently or not, also we note the big disadvantages in Sobel filtering is very sensitive to noise also the accuracy is low for object detection phase, they also used the median filter as we know, the cost of this filter is usually too high, also it is not comprehensive and does not give good results, as for the advantages, firstly they used more than one method for detection the moving objects which are “*Optical flow and Background subtraction”* and cover those topics in several respects and this is was a good idea to offer more robustness ,also they covered many points like low costing, rapid accurately detect, view same place from various angles, the possibility of night vision, they used Sobel filtering and median filtering to reduce the noise.

- The worker team on this paper [7] were discussed and clarified many important things as we shall see. Detecting and tracking moving objects might be a crucial problem due to some blockages that hinder computers from identifying active objects. Therefore, there should be some developed automatic tools to deal with this problem. In this context, till now, many different approaches and algorithms have been carried out to deal with this identifying problem, in order to develop new tools with the aim of overcoming this problematic situation as well. However, a new model called "*Gaussian Mixture Model (GMM)"* comes into prominence among them, with its holistic approach to track moving objects. This model catches attention thanks to its speed in detecting moving objects and its success in dealing with complicated situations.

As we know, the main goal of all the surveillance systems is to identify repeated objects. These surveillance systems are essential for areas requiring high security systems, in today's world, such as banks, roads, official buildings, etc. In other words, advanced detection systems should be applied in order to track the objects in dynamic milieus. In this respect, separating objects and background subtraction images is very important. Thus, this model based on "*Gaussian Mixture Model (GMM)"* proposes a holistic subtraction technique to separate the background subtraction images. All in all, this model, thanks to its subtraction technique and algorithm aims at tracking active objects by dealing with background shadow problem.

As maintained above, there are many approaches that have already worked on the issue of detecting and tracking objects [24]. Even though these approaches have many advantages in terms of tracking objects in a fast way, the main deficiency of these works is that those are not able to determine and to track broken background pixels and to solve the shadow and noise problem. Thus, taking a correct result started to be more difficult. However, of course, some following works made some contributions in terms of reducing processing and recognition rate. Yet these works also fell short in determining and tracking problematic foreground and background objects, calculating video frames and solving shadow and noise problem. Thus, a huge necessity emerged for understanding the behaviour of problematic moving objects to overcome these deficiencies by a new holistic approach.

Therefore, first and foremost, this model based on "*Gaussian Mixture Model (*GMM)", will create the video frames to detect Foreground of the object also pixels of background, Secondly, solve some problems like remove problems of shadow and this will by using HSV model, then they move to pre-processing phase that to decreasing the noise in the image, after that they focus on separating active objects and store it's features in queue intended for feature extraction for compare it with the object that have new moving.

For the results of the proposed system, they applied GMM for algorithm of background subtraction, here in this technique every pixel is represented as a mix that is the amount of K.

Where,

Xt is the density of the pixel in time t from pixel, after that store it {X1,X2,......Xt} also the pixel is represented by use mixture K Gaussian distribution, and expresses probabilities of monitoring for density the pixel in time t by:

$$\mathrm{P}(X_{t)} = \sum_{i=1}^{k} \omega i, tn(Xt, \mu i, t, \sum i, t) \tag{2.2}$$

Also,

ωi indicates to amount of Gaussian groups to models pixel of history.

t is a factor of weight connect with collection of i in the time t.

n is Gaussian pdf.

μ, i, Σi, t represents average and variance array of Gaussian collection.

Every Gaussian ranked depending on $\frac{\omega}{\sigma}$.

P is the estimated total weights of the ranked Gaussian until threshold, which reaches TH=0.25, against of K distributions are checked for every pixel to get the matching, range of pixel value is 2.5 times of standard deviation.

As described above, one of the task which this model espouse, is to solve the shadow problem that is an important part to foreground detection which other approaches has not solved before. In this respect, the model uses HSV colour space model, here HSV values is obtained by existing frame I(x,y) and frame of background such as following:

$$V=\frac{Iv(x,y)}{Bv(x,y)} \tag{2.3}$$

$$S=Is_{(x,y)} - Bs_{(x,y)} \tag{2.4}$$

$$H=|Ih(x,y - Bh(x,y))| \tag{2.5}$$

$$S(x,y) = \begin{cases} 0, \text{if } (\alpha < v < \beta) \wedge (s < Ts) \wedge (h < Th) \\ 1, \text{ otherwise} \end{cases} \tag{2,6}$$

In the next phase, the model implements a process to solve noise and the mismatch among shapes. The model of post-processing, by morphological filtering, where uses a restructuring technique to determine the number of pixels which have been added (dilation) or removed (erosion) especially for the boundaries of object. Afterwards, the extraction phase comes before us. In this phase (Feature extraction), the algorithm used in extracting feature gives us important information about video frames and scenes such as a centroid where detect object It is done by represent the object centroid in regular form about borders of object.

As for the tracking process, the approach of this model locates the important objects throughout their appearance in the video that is given as a sequence. Then, the tracking system applied by this model gives possible positions of the problematic objects during the video and this is provided by the "*Kalman Filter"*. This filter shortly provides a robust estimation about the possible positions of the track within every single frame in the video and reduces the difficulty of calculation; the proposed system has been applied on sequence of images with size 640 * 360 for videos of indoor and other outdoor.

As mentioned before, for many computer based application such as activity and identity recognition, traffic observation and other things related to security concerns, detecting and tracking moving and active objects are very crucial. This model, in this regard, offers an advanced and holistic approach to detect and track this objects, compared to the works carried out before. Simply put, this model, on the basis of "*Gaussian Mixture Model (GMM*)", uses the background subtraction they used. Beside this, this model provides a faster method in terms of detecting video frames and scenes, thanks the algorithm used during detection process. The experimental consequences of this model also verify that this model's results are more accurate and more useful in terms of performance compared to the others. Additionally, this approach is considered as more robust in changing milieus and times. As for the segmentation different kind of moving objects, it is seen that this model offers more certain results to separate these objects and solutions to deal with the problematic moving objects as well. Finally, this model creates a lesser difficulty to calculate video frames, this model (*GMM*) has some disadvantages such as some issues that related with numeral side require dealing with the process of equalization and applying it in order to eliminate the differences also, this will lead to an increase in cost.

2.2.2 Optical Flow:

Optical flow is a method depends totally dependent on the distribution of velocities, the objects that existing in the image, usually it is used to describe the feature in the images, this method is very effective but is still sensitive to noise and this requires a special equipment for the applications which works in real time**.**

- In many papers It was discussed subject of the Detecting anomalous human behavior , In this current system there are some of abnormal behavior activities are not detect it like punching and pushing, the intelligent system in this work is proposed to detect most of aggressive human behavior. The researchers on this paper [8] worked on two specific concepts which are HOG (*Histogram of Oriented gradient*) where indicate to shape based feature and HOOF (*Histogram of oriented optical flow*) where indicate to motion features Then, with regard to modelling method they are used the "*support vector machine* (SVM) "to classify the aggressive events from normal events also they are used a benchmark dataset which is UT-interaction dataset.

For the proposed system the architecture is input the data (video) which has been get it from a camera feed, extract the shape feature by use HOG technique where works to detect the shape of the local object in the framework where can be measure it by two things which are edge directions or the intensity gradients, and can be accomplished by division the current image to 16*16 blocks then each block Composed of 2*2 cells each cell has size 8*8, The Gaussian window size sigma value is equal t. Also touched researchers here to explain HOOF and how it is calculated ,this occurs by computed the optical flow for each frame in video stream and ignore each flow vector based on the primary angle from horizontal axis and weighted according to its size, they adopt the optical flow method according to collect two things which are a data term with a spatial term, where data term depends on the fixity of the image feature and the expected value of the difference for the flow based on the spatial term, the Optical flow has been defined by the following formula:

$$\text{E}=\iint [\left(I_{xu}+I_{yv}+I_{t)}\right)^{2}+\propto^{2}\ (\nabla_{u}\ \|^{2}+\|\ \nabla_{v}\ \|^{2})]dxdy \qquad (2.7)$$

Where I_x, I_y are subsumed from image intensity over the x and y ,u are the horizontal of Optical Flow ,v is the vertical of *Optical Flow*, t is time ,α is the weight of the regularization term, also the Horn-Schunck optical flow method has been applied to get all features with low level movement for each pixel in the frames.

As I mentioned before they used SVM model to classify the aggressive and normal activities, here they have 30 normal file and 29 abnormal file from the feature, which is extracted after that the descriptor feature is formed, the final thing is set a label and trained with SVM.

Final step in this proposed system is alert system where if any aggressive activity is found then the alarm will generate the system will return to first step that is monitoring the system.

This work was good when it used the optical flow algorithm, they provide a study to classify normal and aggressive activities by using HOG and HOOF, also they adopt a discriminative model which is SVM, but there are some weak points which are: in this current system in pre-processing phase there is an important point namely background noise, the researchers here did not cover this main point completely also the proposed system can’t detect many activities just limited activities.

- The aim of this study [9] is to find the events of unusual group in the stream in video, which is the most common and challenging function in computer vision, so we suggested a way which based on the describe of form and rating of approach to handling this problem, that way is start by apply the *optical flow* method then follow by HOFO descriptor descent which uses for detecting the anomaly activity in queues ,this way can be valuation by using descriptor image and available public dataset, as a feature of HOFO descriptor is to find unusual movement, “Support Vector Machine (SVM)” used for classification program. There was been install surveillance cameras for safe reasons due to the development of technology and the improvement with life style, the revelation of the behaviour of unusual group is a real problem research in the computer vision and system of surveillance. Computer function is a field with a method that used to own, handling, analysis and comprehension image. Analysing events that preternatural in external environment is the most known functions in the computer function.

In this research, the suggested way is to utilize and to identify a set of unusual movement in the video is histogram optical stream introduction (HOFO) which consolidated with some class bolster vector machine (SVM). HOFO descriptor is utilized as highlight to recognize suspicious moving lines and furthermore models a halfway image. SVM is utilized for characterization strategy. The suggested system is intended to recognize the unusual happening which has happen in the video scenes, first of all the user will choose the video then detects the abnormal activities in the video, from the PETS dataset or UMN dataset the user will choose the video which includes the video with normal or abnormal activities, with a specific end goal to distinguish gather exercises in an unsupervised way for outside environment, the processing of video is finished.

In Computation of *Optical Flow* they said “there are many frames which are exist in video so we have to extract some of these frames from it”, when all the frames have been extracted, the next step will be calculate with the optical flow for every frame in the input video, these frames are sequences of videos which have a little time between all previous frames, optical flow uses especially for objects which are moving, in this procedure the initial phase incorporates calculation of the optical flow for the features but not for all, this just for grey scale, then it will apply on the sequential frames which have been acquired from the video, so the conclusion of this procedure is a movement vector that perform to the optical flow, also this method can be made by focusing on paring frames to the focuses in the following frame, this kind of points are finished by contrasting the densities of those points in a window that is given, we can make the pairing points by contrasting density of point with another one which has the littlest difference in density, then the matched between points is done , the speed is computed by range of points which are moved. The HOFO descriptor is a feature which utilized to distinguish anomalous moving; in here every frame is split into many blocks, after that HOFO is calculated every block, also it is calculated on a thick grids interlaced block then collects the Histogram of every block to obtain a vector with global nature, a probable vote of every pixel is computed, it is depends on the components of optical flow, after that all these votes are collected into orientation boxes which done over local places districts.

There are two types of techniques that exist in machine learning, these techniques are supervised and unsupervised learning, for first type is a collection of data under training which utilized and takes as inputs then analysis it, the second type is utilized for rating, they use SVM technique to recognize specific event this technique is used for classification, so they acquire a support vector based on every frame which will rated also the event will be found, so when the result of that rating will detected, it will alert if any suspicious event will be found.

There are two dataset PETS 2009 and UMN dataset which are used to examine this method, the datasets have aims which are discover reason and happening such as walk, run,…etc. PETS 2009 datasets have different video scenes, some are normal, other are abnormal events, Initially they prepare these video sequences and train it to identify the unusual events then they test the proposed system by focus on all different snapshots video and take them as inputs, the normal event is presented as a set of people walking without making any suspicious activity, as for the abnormal events are presented as a set of people which are running with quick activity. UMN dataset have different video scenes, like yard, square, indoor, the aim of this paper is the outside environment events so they can implementation this suggested approach on all those scenes, this kind is utilized for preparing and test this system on these kind of activities that include a set of people which are moving with different direction, as for abnormal scenes that are include a set of people that are running, the result of detecting abnormal events that depends on HOFO descriptor that give us good results. The basic idea in this paper is the results of this proposed system are based on distribution the histogram which is related with optical flow and his features also on HOFO and SVM method that are for classification, where SVM technique is a global behaviour so it is a robust and effective in outside.

They also used a confusion matrix, composed of set of information which are represent activities between actual and predicted events, also the diagonal in the matrix are a set of a correctly sample and other Incorrect sample which has error cases, where the Diagonal of the matrix is composed of activities which are walk, move and Run, after that they take the video which has these different activities and train it. As following:

Accuracy=98.99%

20	0	0
1	32	0
0	0	46

Confusion matrix

As a result the proposed system on PETS 2009 and UMN dataset, also on confusion matrix this show us an effective results and a high performance. But there some weak points like this system is apply just outdoor environment also just on some activities, so this system is not comprehensive system, but it covers just some of these possible events, also they used SVM method and there are some disadvantages in it which are the results are not transparent as well as the ratios of financial are very high.

2.2.3 Frame-To-Frame Difference:

This method depends on the difference between the frames that are sequential in sequential images and this is in order to detect the regions that corresponding for objects especially moving one like vehicles or people [10, 11]. Where they play the threshold value a big role in this method, this method is really an efficient method in dynamic environments. This method is considered as very close to the "*Background subtraction method*".

- The researchers in this paper [14] worked on "*Mean Feature Point Matching (MFPM) algorithm"* for detecting the unusual events also the "*Speeded-Up Robust*

Features (SURF) method" is proposed for extraction the features. Where MFPM algorithm works by comparing feature points from the input image with the feature points from trained dataset.

The proposed system here depends on using cell phones as an unusual event in private regions The system works by taking the material of videos from the camera and use it as input, then detects the unusual events by using trained dataset. After that is applied the SURF technique on a huge set of images which are include unusual events ,this usage happening in different modes to provide larger space to train the system on different segments of cell phone usage. We note the system detected feature points in sample images and extract feature the descriptors in the important points.

This system has been designed based on use 150 important feature points in each image from sample images, then computed the mean feature point and stored for more processing, after that apply MFPM algorithm on input frames to detect the unusual events. As for feature extraction Speeded-Up Robust Features (SURF) method is proposed to detect the blob features, the SURF method used Hessian matrix to feature extraction knowing that Hessian matrix is a second derivative matrix, as for the feature description it is used SURF algorithm which uses the wavelet responses in horizontal and vertical directions and taking size of neighbourhood M*N about the main point, also divided it to subareas where for each subarea is computed by vector V:

$$V= (\Sigma dx, \Sigma dy, \Sigma|dx|, \Sigma|dy|) \tag{2.8}$$

The total of dx ,|dx| calculates separately for $dv < 0$ and $dv \geq 0$,also the total of dy and |dy| are divided depending on the sign on dx after that doubles numbers of feature points, those feature points will be extracted from the sample images and stored it in M*N matrix this matrix it will be converted to single dimension array for more processing and this by calculating the mean feature points, this mean feature point is average value in every column of M*N feature matrix , the final results it will be to detect the unusual events.

In MFPM algorithm is working by taking the inputs as forms of video then converting the video to group of image frames also cancel the similar frames and convert the input image frames to grayscale mode for more processing.

1. Input the sequences of frames (∑.)

2. Processing the frames.
3. Put F_0 as a first frame.
4. Move from first frame F_0 to last frameF_i, i<N_f.
5. All frames F_i are fragmented to quadrantsQ_i.
6. These quadrants are compared with the MFP of the object that exist in trained dataset.
7. Plot all inputs that contact with Q_i from F_i and Obj_i from T_i in the matrix $M[Q_i, \mathrm{Obj}_i]$.
8. Matrix of features are calculated by: $\begin{cases} 0, \text{if } Q_i \text{ and } \mathrm{Obj}_i \text{ is matched} \\ \quad 1, \text{ otherwise} \end{cases}$
9. C=C+1 when the value of $M[Q_i, \mathrm{Obj}_i]$ is 1.
10. The ratio of matching are calculated by R= C/N_f .
11. Return the matrix of features matching $M[Q_i, \mathrm{Obj}_i]$.

The results of proposed system shows us that this system is capable on extract the features of interested area from huge image samples, where they used the captured videos as input also from the results can note this system capable to detect each usage of the cell phone in each frame , also the researchers showed us a collection of images to detect by usage of cell phone and in any quadrant from the image this happened, and how the system move from quadrant to other in the sequence when find the object in any quadrant ,and therefore it will reduce from the time complexity.

This system has a high efficiency in detecting unusual events which is here cell phone usage in private places like planes by using the mean feature points matching method, The experimental results show us the efficiency of the system by Implemented it on the input videos, but there is more cases we can add it to the system to test if it able to work with high performance or not, because just one case which is detecting of usage the cell phone not enough, the results also not discussed enough to figure out what the system was already working efficiently where they discussed only some cases without covering the subject by significant number of results to prove the efficiency of the system.

2.3 Object classification:

The objects classification [23] is one of the most important phases in video surveillance system as the main purpose is to extract the corresponding area from the object where the extraction occurs for all moving blobs[12,13]. In today's world, to understand and to recognize human motions in videos is a very popular research area. For years, lots of works and researches has been carried out the human recognition process in videos. For example, some works using two-dimensional video has focused on "*background subtraction*" along with "*Gaussian Mixture Model*" to carry out segmentation process. Moreover, some other researches use the "*spatio-temporal bag of features*" for determining action. And in order to classify, they use "*a non-linear support vector machine*". There are two main sections for classification or recognition phase that are "*shape-based classification*" and "*motion-based classification*".

2.3.1 Shape-based classification:

In the classification based-shape, a various characterization of shape information is provided such as a box, blob area…etc. They used them as standard patterns in classification subjects to classify all objects that are moving in the movie such as "R.T. Collins and et al". [15] Where they divided the objects that are moving to some parts by using "neural network classifier", the input features were a mix from images that contains different scenes with different parameters to classify the shape according to those parameters also they applied the classification on all frames and each moving blob inside the frames, the classification results were saved by diagram, in order to be the results more accurate the temporal consistency are taken into consideration.

- As known, the classification process of objects has several phases. In this paper[16], the writers discussed the subject and divided it into three phases, the first phase, as expected, they are discussed detect the object and recognition correctly and then to determine the features of the object detected. As for the third phase, they tried to reduce the dimension of the object and finally classify it. However, in this respect, they presented a study about a geometric also appearance feature ($\in R \approx 25000$) especially for a system surveillance in outside scenes, also they focuses on elaborating on the various features of the objects and uniting them. Therefore, this study's contribution in the literature, creating an object detection with high accuracy and comparing the different reduction systems for dimension of the objects and comparing the various classification algorithm to be used.

Till now, lots of researches on the object classification have been carried out. These researches have mostly focused on comparing the methods and algorithms used for the classification of objects. These works have improved some features of the detection systems such as solving intensity, repeatability, and colour problems. However, they have fallen short of solving the occlusions on the background and cluttered points in the video. And, as known, to fix these problems, highly advanced and descriptive classification process have to be developed, where they talked about object classification in images also in videos and the algorithms which are used for this such as "Maximally Stable Extremal Regions (MSER) ,The Difference of Gaussian (DoG), Support Vector Machine (SVM)".

The basic goal of the surveillance systems is to detect motion and activities. For this reason, in this respect, the most important thing is to separate the background images from the moving object. It can be classified, the algorithms of segmentation to four phases which are: background subtraction(which is considered the most important approach), segmentation the density of motion, also segmentation the video, finally specific object detector, In this system is build a background model by use non-parametric(KDE) ,by using this model each pixel it will classifies to background or foreground. As result for apply previous model get a set of foreground areas that connected together, and to this end, this algorithm carry out this process in an accurate way by color distribution in subareas (Blobs) and geometry for the first object extraction and track it.

In the next extraction phase, the process is carried out to measure the effectiveness of some of descriptors and to compare various algorithms extracting the objects accurately. In this phase, "*Histogram of Oriented Gradients (HOG)"* is used to detect and classify objects, this algorithm builds a histogram for values that are separate from gradient orientations to detect the objects, also they proposed "*Luminance Symmetry"* in this study where used this technique to measurement the brightness symmetry of the objects, they calculate "*Luminance Symmetry*" about axis by:

$$L_{sym}=\frac{1}{C}\frac{2}{w}\sqrt{\sum_{i=1}^{h}(\sum_{j=1}^{\frac{w}{2}}I(i,j).B(i,j)-\sum_{j=(w)/(2+1)}^{w}I(i,j).B(i,j))\text{^}2} \tag{2.9}$$

As for Central Moments also translation and finaly rotation they extracted 7 Hu moments this applies on image of object for ADI where this object is the absolute value of the differential in image and this before thresholding by other word after background subtraction oriented to a box about the object, they also discussed Angular radial transform (ART), this technique is useful in compacted and capturing all regions that related and non-related with each other

they extracted ART descriptors with the standard configuration

$nAngle = 12$, $nRadius = 6$ which gives in total $6*12-1=72$ features.

In the Cumulants phase, they used 3 textural characteristics which are: the density of mean value (E[X`])`, the density of histograms for Standard deviation (E[(X-μ)^2]`), the density of histogram for Skewness ($\frac{E[(X-\mu)^3]}{E[(X-\mu)^2])\text{^}3/2}$ as for Horizontal and Vertical Projection has been processed by:

$$HP_{I1+}=\sum_i B'(i,j), VP_{+,j}=\sum_j B'(i,j) \tag{2.10}$$

Where:

HPi, +: is Horizontal Projection, *V P+, and j:* is Vertical Projection (for every bin).

I: is the rows, *j:* is the columns.

For *Morphological Features* they extract some features which were the *Anthropometry* $A_{th}=\frac{H}{P}$ that is a static percentage to body of the human, *CompactnessCmpct*=$\frac{Ar}{P\text{^}2}$ for measuring the shape difficulty, p is the circumference for shape of the object, also *Aspect ratio* *AR*=$\frac{W}{H}$, w is width of the square of the object, h is high of square of the object, and *Solidity* $SD=\frac{Ar}{ArcH}$ for measure the parts that curved inward from the shape, Ar is the circumference for area of the object , ArCH is that part from the cambered Hull which include the object.

To reduce the dimension is another important step for this model. In this regard, two different reduction techniques are used for this step. Firstly, in feature transform phase, the dimensions of the new place(linearly - nonlinear) is the transformation of the original place. However, in the future selection phase, the low dimensional place's new dimensions is the sub-group of the high dimensional place of the original place.

In this system they used PCA as a method for feature transform which is unsupervised and briefly, thanks to this feature selection phase, the high dimensional place can be turned into a low-dimensional place. Also they used a specific method for a feature selection which is "*Entropy Based Discretization* (EBD) " where they compared between three features which are the original features that extracted also feature selection and feature selection.

With the aim of *traning*, the model firstly extracted some various motion images from "*VIRAT dataset"* such as human, vehicle, car, bicycle and object to demonstrate the accuracy of recognition. Also, they imposed some constrains which were: Detection Percentage(Dp > 30%) where has been computed by: $Dp_i = 100 * \frac{C_{arcai}}{BB_{arcai}}$, Overlapping Percentage(*OP* < 10%) which computed by: $Op_i = 100 * \frac{\sum_{i \neq j} \cap (BB_i, BB_j) arca}{BB_{iarca}}$, motion and object instance constraints, where BB is bounding box area. During sample taking process, the *object classification* is carried out by the techniques that are SVM and AdaBoost by comparing the results of both. For classification the multi objects they used C-SVM, AdaBoost technique, in this context, is successful to combine inadequate results to achieve the strongest classifier.

As for the experimental results of this model, *apperance based classification* on the basis of HOG features they used VIRAT dataset, also they used C-SVM (80% − 20% training test) where they achieved 71.4%, PCA based SVM classification here they used a set of features with 30-D where divide to 2 parts(80% training, 20% testing) the results were really good with accuracy 89.9%, ,for *feature selection based SVM classification* where they used 142 features of the total 25,000 features, also here they have been applied C-SVM performance on (90% − 10%) ,the best result was (accuracy = 92.3%), Finally, in the "*feature selection based AdaBoost classification"*, the most crucial work is to detect weak classifiers and increase the number of them in order to create more solid and accurate results by combining them with the help of a proper formula. And as a result of four experiments on the 80% − 20% training-test, even though it is seen that appearance based features did not perform well, geometric features performs importantly better in video surveillance systems.

For low resolution problem with detecting objects in the Surveillance Systems ,use only HOG feature is inadequate to recognize the objects. However, they said if we combine it with geometric features we can achieve high accuracy for recognition in the surveillance systems. For this, this model has used geometrical features named luminance symmetry, central moments, ART moments as well as HOG features. And afterwards, with the help of feature selection, the combination of features is created for the object recognition. And lastly, SVM and AdaBoost classification techniques are applied for adequate and accurate recognition objects by combining the classifiers in an effective way.

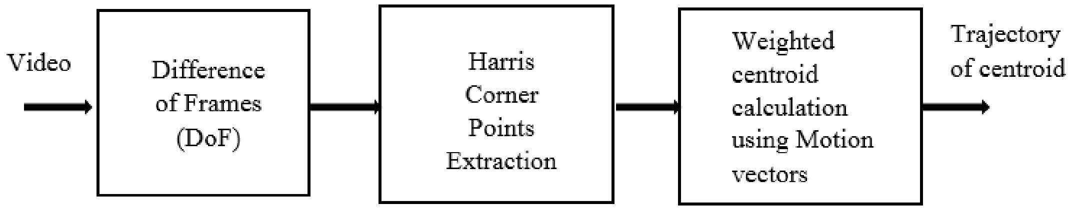
this study is a comprehensive study where the researchers discussed a set of techniques and compared between the results where were excellent results, But the problem of classification of small objects still exist so they should intensification their work and focus on this point to get the most accurate results.

2.3.2 Motion-based classification:

The concept of this method is based on features of the motion objects that features must be unique to be able classify the different objects ,the classification-based motion [17,18,19] is usually used to distinguish between objects that are solid like vehicles and other non-solid objects such as human, some studies have relied on the temporal similarity of objects that moving.

- Methods of human action recognition are having many difficulties due to the complexity of the real-time applications because it contains numerous parameters. In this study[17], a different approach is explained which is based on the idea of using "*Difference of Frames (DoF)*" by extract Harris corner points from the motion vectors that used to identify the centroid locations that obtained in each frame and use all those corner points as weights. With the sequence formed by the motion, where used the sequence of quantitative orientations as "*temporal features*" and this is used to classification the different actions trajectory that movement of centroid creates which use the models of probabilities state to make the decision. Training and testing of state models are done with various benchmark datasets like " KTH dataset, Weizmann dataset and UIUC complex activity dataset". This approach worked well with identifying complex action classes like walking or running by being applied to wide range of intra-class variations.

Latest technological developments that enabled large video contents caused the problem of classification of the actions performed in them. A typical action recognition approach consists of two steps that are feature extraction and action classification. A typical action recognition approach consists of two steps: feature extraction and action classification that use a suitable machine learning tool. Challenges like motion performance variations, moving backgrounds, changing viewpoints or low resolution caused the automatic human motion recognition from videos to be very difficult. There are many studies that are based on "silhouettes extraction",also " robustness to change in color and contrast of the video" however recognition processes' complexity should be lowered by accurate person localization algorithms. Other methods which is based on optical flow that avoid use of a background segmentation process via encoding the information of motion between two consecutive frames in the video, but unfortunately these methods are extremely sensitive to the noise and occlusion and are computationally intensive. However, the usage of both representations was limited which are the representations that based on flow and the other representations which based on silhouette that try to encoding of huge amount of visual observation, local feature extraction methods are preferred over these two which are based on identification of areas that faces important variations in spatial and temporal directions. "Laptev and Linderberg" detected spatio-temporal interest points using 3D Harris corner detection algorithm and "Bregonzio" put Gabor filtering on DoF in the action sequece to detect the important points that base on the way of modeling the temporal content. Methods of action classification are split into two differente categories according to how temporal content is modeled. First is, direct classification where the features of temporal block is treats as one entity without an obvious temporal content modelling. Second is the temporal state models which are based on the temporal variations in action. (HMM) is a versatile probabilistic technique, HMMs fail to model parallel actions in the video due to their being sequential models of action. "*Dynamic Bayesian Networks (DBN)"* handle this restriction well. There are works where DBN is used to model interactions between two persons or person and objects. This paper aims to simplify present methods using trajectories. The more feature is effective between all other features Motion trajectories due to their compactness like silhouettes. The reason is that the motion is represented through a sequence of pixel locations for some distinction in next frames, the result of this work is production of a trajectory which clarifies the point of mass center for the motion between the next two frames, we can see the generating trajectories by the block diagram as below:



Trajectory Extraction.

Where it is assumed that the background is static, the action is performed by a single person and no change occurs in camera position, also the change in the action in each frame is clarifies the action of recognition. The points that extracted by "fixed Harris corner technique" are provide the data to explain the shape of interest in an image, as well as a close concept of the region encoding process that obtained from the action change in DoFs.

For “*Weighted Centroid Calculation*”: The corner points are clarified as a set of locations where find a big differences in all neighbourhood directions, in every image from DoF images the spatial distribution for the corner points are represent the action of transition information where the centre of this distribution is called centroid. Some parts have higher motion than others when an action is performed and these are preferred to track the key location due to their high accuracy and reduced false positives for this reason, centroids are calculated by use the motion vectors. They defined the motion vector as 2D vector which presents the offset between the (X, Y) -axes of the block “coordinates” in reference frame and (X, Y) -axes for same block “coordinates” but in the following frame, the present frame divides to a matrix consisting of macro blocks which contain a set of corner points and these points are compares with "the corresponding blocks" also with other blocks In the area to be searched, the changes in the location of corner points give us the offset vector. This way, computational work is reduced by performing the estimation only in definite points in DoF. Then, every corner point is connected to magnitude of motion vector as following:

$$(x_c,y_c)= (\frac{\sum_{i=1}^{n} d_i x_i}{\sum_{i=1}^{n} d_i}, \frac{\sum_{i=1}^{n} d_i y_i}{\sum_{i=1}^{n} d_i}) \tag{2.11}$$

Where:

$(x_1, y_1),(x_2, y_2),\ldots\ldots.(x_n, y_n)$are “key points”.

$d_1,d_2 \ldots\ldots\ldots\ldots. d_n$ are “corresponding magnitude of motion vector”.

Where the threshold value is 0.001 that to measure the Harris corner and put the value (2 pixels) for parameter P which Indicates to search area value.

To differentiate between similar motions like walking and running, trajectory orientation at segment level is used as a feature in which tangent is calculated by displacement in vertical axis being divided by horizontal displacement. Also they Indicates to Segment Orientation by:

$$\text{Segmentorientation}=\tan^{-1}\frac{x_2-x_1}{y_2-y_1} \tag{2.12}$$

Where:$(x_1, y_1),(x_2, y_2)$ are the centroid values for sequential frames.

As for "*Intra-class variations*”, Bin assignment concept is employed to classify different styles of the same action together which quantizes trajectory segment orientations. This is done by division of range of orientation into eight equal bins that have a bin number. Then the trajectory of the motion is represented with bins to be processed later. Since the bin assignments are "time-series" also, we know that this is a random data in nature, as that "probabilistic state models" considered as better than the deterministic ones this is with regard to a model of a temporal variations. Where we can learn all of the "*state transition probabilities*" and "*state occurrence probabilities*" through a training stage by using a huge of samples specialized for training. They calculated the "state transitional probabilities" in "state transition matrix" by:

$$a_{ij} =port_{(}t_i/t_j) =\frac{count_{(t_i,t_j)}}{count_{(t_i)}} \tag{2.13}$$

They assumed that actions have an equal probable, so "maximum likelihood (ML) classification" is used with state models of twko action class to identify the action in test video. The "sequence of bin assignments", "corresponding probability conditioned" is computed by:

$$P_{(obin/action)} = \prod_{i=1}^{N-1} B(i)A(i+1) \quad (2.14)$$

Where:

A: state transition matrix for an action class.

B: corresponding state occurrence matrix.

N: length of bin sequence.

$$P_{\left(\frac{Walk}{Obin}\right)} = \frac{P_{(Obin/Walk)}P_{(Walk)}}{P_{(Obin/Walk)}P_{(Walk)}+P_{(Obin/Run)}P_{(Run)}}, P_{\left(\frac{Run}{Obin}\right)} \frac{P_{(Obin/Run)}P_{(Run)}}{P_{(Obin/Walk)}P_{(Walk)}+P_{(Obin/Run)}P_{(Run)}} \quad (2.15)$$

The proposed approach is based on walking and running actions classification using "weizmann" and "UIUC complex activity"datasets. This approach's accuracy rate is high. Accuracy reported on "wiezmann dataset" for walking action is 100% whereas on "UIUC dataset" is 85%. For running action the classification accuracy on"Weizmann dataset"is 100% and that on "UIUC dataset" is 80%.
In this study discovers computationally simple action recognition that can generalize large set of intra-class and anthropogenic variations . It reduces the human involvement also. To extend this approach to detection of parallel actions or to definition of new variant, trajectories at limb level along with the trajectory extracted in this approach can be used as features. Also, multiple trajectories can be handled by designing a good framework.

This work was good because they worked to reduce the complex arithmetic operations for action recognition, also results are indicated to a high accuracy also they used a less involvement for human and this makes work much easier, but here the authors were rely entirely on just two data sets for the whole study and this does not made trust it high for the results that have been obtained.So for this reason we can use other data sets to be the results more comprehensive.

- In This approach [18] shows a perfect performance along with "*KTH action dataset*". Beside, some other researches use "*Spatio-temporal techniques*" to reach levels of abstraction that basis to "*Scale Invariant Feature Transform trajectories* *descriptor (SIFT)*" In these studies, non-linear support vector machine is utilized to classify the actions especially for HOHA and LSCOM datasets or human movement in other words. Plus, human movement is showed by temporal templates with the help of Motion Energy Images (MEI) also the descriptors of Motion History Images (MHI). However, most of the previous works were restricted to determine the simplified human motions. In this respect, this study is eager to determine interest point based on “Spatio-Temporal Interest Point (STIP)”. Afterwards, the research aims at “Extracting Histogram of Oriented Gradient (HOG)” and “Histogram Optical Flow (HOF)” descriptors to show both appearance and human motion. Consequently, the study uses the "Support Vector Machine (SVM) "along with "Bag of Word (BOW)" of the features that are the most interest points and apply the method on “Multimodal Human Action Dataset (UTD-MHAD) ".

The method that used in this paper is summed up as following: the recognition of human movement is process consisting of two steps named training and testing respectively. This proposed is using the "BOW" from "HOG" also "HOF" to presenting the actions, in the training part, we match SVM to the model of collect action with SVM parameters. Then utilize the outputs alongside with test video to obtain the exactness of recognition.

In order to show the human action, the study firstly extracts the STIP, HOG and HOF. As we know, HOG and HOF are two approaches that are most popular in this field. Then we begin to use the space-time interest point to identify the sequences of the video. Moreover, as we mentioned before HOG and HOF are extracted to show every interest point by using "Harris interest point detector". The spatial HOG descriptor shows the appearance and shape properties whereas HOF identifies the local motion. That is to say, HOG identifies the shape and appearance on the basis of distribution of the intensity gradient. They divide the image into cells in small sizes and calculate the histogram of gradient direction and gradient orientations by means of a computer. The calculation of the gradient (G) and orientation (O) as seen in the formula below:

$$G = \sqrt{G_x^2 + G_y^2} \quad , \quad O = \arctan\left(\frac{G_x}{G_y}\right) \qquad (2.16)$$

In this formula G_x and G_y represent the horizontal derivatives and vertical derivatives of every interest point.

As for HOF descriptor, it shows the local motion of the interest point and it carries out the process based on the "luminosity conservation hypothesis" as seen in the equation below:

$$\mathrm{I}_{(x,y,t)} = \mathrm{I}_{(x+dx,y+dy,t+1)} \tag{2.17}$$

In this formula, “I” represents the “intensity luminous”.

They calculated the motion of vector for every cell by means of a computer and build the orientation also "magnitude representation". Every flow vector is voting to bin and this based on the angle also on the magnitude with the weight. The sequential of all histograms of "*Optical Flow*" is display "*HOF descriptor*”. Afterwards, they implement "*K-mean quantification algorithm*" to those merged extract descriptors with the aim of build the “*Bag of Word*”. While BOW shows the properties by utilizing a visual vocabulary, the establishment of codebooks is carried out thanks to HOG and HOF features. Finally, within the “*Spatio-temporal bag of features”*, the video series are shown by the histogram of visual word events over the volume of space-time.

This study uses the "*Support Vector Machine (SVM*)" with the extracted properties that analyze and identify the human movement. SVM is limited to one kind from classifications issues that is the binary classification issue; however, there is an extension that is carried out for multi-class classification later. In multi-class classification, two main approaches are offered to merge various binary classifiers and react all the classes in an immediate way. While merging some binary classifiers term, they see "one-against-one" and "one-against-all" techniques. As for first strategy, one builds one SVM for every pair classes and the other strategy produces one SVM for every class that is utilized to differentiate samples between one classes from all of the classes actually. From the max of SVMs outputs, they can classification data that are unknown. After comparing these two approaches, the writers see that one-against-one approach is much more proper for in training samples. For that, they prefer to use one-against-one approach in this study as well.

The multimodal dataset used in the process encloses "RGB video", "skeleton joint positions", "depth video" and "inertial signals data" besides that, a "Kinect camera" is carried out to gather the three data: "skeleton, color and deep images". As for the data, number four of acceleration, "magnetic strength and angular velocity" are provided by the inertial sensor signals. The database are contains 27 actions and these actions are experienced by four men and four women. These actions are repeated for 4 times by every subject. All in all, total 861 data series emerge after we remove three sequences that has been corrupted. In order to prove our study's usefulness for identifying human action, they experience UTD-MHAD dataset "University of Texas at Dallas Multimodal Human Action Dataset".

Due to this algorithm, they use approximately 60% of the color succession video in the training part and the remaining 40% is used for the testing part. The rate of recognition changes from one action to another. The authors see that the overall recognition rate of their method obtained through UTD-MHAD is 70.37%. With the aim of developing this approach, they carried out a comparison process. As a result of the comparison "DMM-CRC" using depth data, also the approach of "STIP-BOW-SVM" using RGB data, they see that their approach is obtained a recognition rate of 67.37% while the first one obtains 66.1%. the main reason of the difference of 1.27% is that their approach uses color sequence video better than the first method using "Kinect camera "date of the "UTD-MHAD dataset". In this study, the authors proposed an approach to catch the human action recognition by using a color video sequence and presented the video action series by a "spatio-temporal bag of features" on the basis of STIP, "HOG and HOF descriptors and BOW". And the process was supported by "the support vector machine (SVM) classifier" to classify the human actions. As for the experimental results of the study, this study based on RGB data perform better than the UTD-MHAD using method. In this study, only the RGB video is used while UTD-MHAD is a multimodal database that contains depth, skeleton joint positions, inertial signal data and depth. Therefore, we can say that the other modalities data or carry out the mix method to develop the exactness that will be studied in the future researches.

We can note that this study used a new approach, where used many methods were mixed to reach the best results in the field of recognition of human actions as well as increase the accuracy of the results obtained previously, but there are some gaps in this study where the results show that the accuracy is not enough and up to 67.37 % This result is not considered convincing and therefore more training on the data should be done to improve this results.

- This current study [18] focuses on real-time human action classification by using original and depth map series and propose a reasonable time obtained by a computer system. In addition, it presents "a new Adaptive Temporal Sampling (ATS)" descriptor that uses the sampling process. Where that descriptor is showed us how it is efficient also saving the time. The researchers observe frames in time based on "the slop of corresponding motion energy curve graph". Afterwards, “the Depth Motion Maps (DMMs)" are calculated. These Depth Motion Maps ensures us to catch the basic and major features one action and provides it in a more discriminative way for the video. Alongside with Depth Static model that describe the features for specific action, this study proposes the new descriptor "ATS-DSM" instead of HOG, extracted the features are used to reach a high efficiency. After that, given that high dimensions and lots of data, “Principal Component Analysis (PCA)” is utilized to make the recognition process is faster. From the results they observe that "ATS-DSM descriptor” accomplishes calculation the real-time and outperform the other methods. According to the method of Adaptive Temporal Sampling, all of the depth sequences is calculated by a computer then three sub-actions are determined from these sequences. These three sub-actions are planned into "three orthogonal Cartesian planes". Afterwards they are merged to build the ATS descriptor. For "Temporal Sampling phase" they can find some algorithms that try to determine human movement based on” bag-of-words and histogram" .Furthermore, these algorithms have a coarse temporal scheme that divides equally the deep image sequences into lower parts by a frame index. That methods constantly have a problems, and these problems include in confused between the actions which have same characteristics, for find a solutions for all those problems the writers suggested put "adaptive temporal sampling method" based on:

$$\mathcal{E}(i)= \sum_{v=1}^{3} \sum_{j=1}^{i-1} sum\left(\left|I_v^{j+1} - I_v^{j}\right|\right) \quad , \quad g(i) = \mathcal{E}(i) - \mathcal{E}(i-1). \tag{2.18}$$

Where:

$\mathcal{E}(i)$ is motion energy, I_v^j is the projected maps.

g(i) is motion energy gradient of I frame.

Also there are some other methods based on pyramid technique are not adequate to handle temporal change as well. Therefore, they propose "an adaptive temporal sampling method" to deal with this problem. A frame’s gradient can give information about its relative variation rate of moving energy. Moreover, this gradient shows the relative moving range between all of the action sequences. In this respect, this offered adaptive temporal sampling method chooses three of the largest gradients to divide the depth sequences. Beside, to balance calculating speed obtained by computer and the exactness of recognition and to form three sub-actions, this method chooses seven frames as well.

As for of phase "Depth Motion Maps (DMMs)" the researchers said calculated it via projecting "the depth map sequences" to 3 "orthogonal Cartesian planes" then computing the difference value for "sequential projected sequences", these maps which are renowned for calculating the exact difference among sequential projected sequences, can identify corresponding moving types in an efficient way. Without threshold, this method calculates the motion energy by computing. In particular, in these maps, every depth sequence frame constitutes three two-dimensional maps which are equal to side or top scenes. As for DMM_v is clarifies by:

$$DMM_v = \sum_{i=1}^{N-1}(|map_{i=1}^{i+1} - map_v^i|) \tag{2.19}$$

Moreover, through these motion maps, we find three sub-actions and three-depth motion maps of the actions are calculated by computer. These "*Depth Motion Maps (DMMs)*" can only catch the dynamic portion of all the actions. Therefore, in order to catch static properties they prefer to use the descriptor named ATS-DSM .They reach" *Depth Static Model(DSM)*" by carrying out two sequential projected sequences alongside with a threshold:

$$DSM = \sum_{M}^{i-1}(|I_v^{i+1} - I_v^i| \leq \phi) \tag{2.20}$$

Where: ϕ indicate to the threshold, and I_v^i indicate to "*the projected map*".

Considering that DSM is reached via all of the depth sequences. This "*Depth Static Model descriptor"* concentrates on the static frames between all of the sequences so that it can obtain the static portion of the action in an efficient way. Moreover, beside this, the model does not utilize HOG descriptor to make the calculation process easier. Instead, the model uses vectorization operator for all the actions. Given that the achievement of real-time human recognition, they use $l_2 - CRC$ classifier to sort the actions in this study. This $l_2 - CRC$ classifier is arranged to reach real-time recognition on the basis of scattered representation method.

In order to test the exactness of this method, the experiments of study are applied on the public field "Microsoft Research (MSR)" that are reached through "*RGB-Depth camera*". Furthermore, the experiments of the current depth-based studies is compared with this method. "The MSR-Action3D dataset" consists of twenty actions and these actions are obtained by ten other subjects as well. The dataset merges the intra-class change since every subject conducts the same action in a different way. Therefore, in order to get accurate comparison, all these actions are divided into 3 sub-sets and then these 3 sub-sets are also divided 3 other parts to conduct the evaluation and comparison. Moreover, by means of 3 tests, the process is carried out. And, during these tests, the first and last 4 frames are removed before constituting DMMv. Also they remove these frames since the subjects are static and not adequate to catch features. Beside, these little body actions hinder us from obtaining the real-body images since those actions occupy large pixel values. Plus, “principal component analysis (PCA)" is carried out to develop the classification calculating. This method differentiates from those second depth-based studies, where the tests show that this exactness in the recognition are over 97% and compared to other studies, this method is superior to that of other studies as a result of many tests carried out. Plus, the tests indicate that ATS-DSM descriptor could make the recognition more accurate. And Cross Subject Test shows that a recognition rate of 93.4% is reached. Furthermore, our study that does not use HOG descriptor is much more efficient compared to other that use it.

They used the confusion matrix in this study especially to apply on "*Cross Subject Test* ".This subset is more intriguing that of "*non-cross-subject*" tests. For that, we can see some reasons for miscalculations. First of all, the performance for same action may vary due to performance of different subjects. And this causes intra-class variation to be larger. Secondly, the classifier makes classification on the basis of re-building errors of various training series. Consequently, the sub-actions chosen from the action sequences may delete other main sub-actions when these actions are too complicated.

This study, the basic idea revolves around depth images, where presented a computed efficient descriptor and ATS-DSM carry out the process. By using both ATM and

DSM they could catch the dynamic and static properties. Afterwards they merge the information as a descriptor to constitute an effective demonstrator for actions of people. The average recognition rate of this method is 96.9% and this means that this method which allows to complete real time human action recognition is superior to other methods, taking into consideration this rate. In this system, we used the previous two types of classification, as we used the shape-based classification to determine the outer shape of the object (human), and we used the motion-based classification by calculated the speed of the moving object within the video.

2.4 Object Tracking

This phase tracking[4,5,6,7] is important for analysing the object motion in video where the object that exists in video is tracked through the sequence of video frames, the system must be able to track the objects as well as the prediction of positions in this system we used Blob analysis technique to analyse the motion of each path this technique is used to predict the different paths for objects which contains in video, also this technique is include high accuracy, good performance and facilitates the complexity in calculations.

Tracking is done by drawing a rectangle around the moving object within the frame based on blob analysis technique [21] as shown in figure (2, 3).

We have used the following code to perform the tracking phase in this system:

```
function pushbutton5_Callback(hObject, eventdata, handles)
% hObject    handle to pushbutton5 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
foregroundDetector = vision.ForegroundDetector('NumGaussians', 3,
...'NumTrainingFrames', 50);

videoReader = vision.VideoFileReader('rpn1.mov');
for i = 1:52
    frame = step(videoReader); % read the next video frame
    foreground = step(foregroundDetector, frame);
end
figure; imshow(frame); title('Video Frame');
figure; imshow(foreground); title('Foreground');
se = strel('disk', 13);
```

```
OpenB1 = imopen(foreground,se);
filteredForeground = imclose(OpenB1,se);
figure; imshow(filteredForeground); title('Clean Foreground');
blobAnalysis = vision.BlobAnalysis('BoundingBoxOutputPort', true,
...
    'AreaOutputPort', false, 'CentroidOutputPort', false, ...
    'MinimumBlobArea', 150);
bbox = step(blobAnalysis, filteredForeground);
result = insertShape(frame, 'Rectangle', bbox, 'Color','red');
num = size(bbox, 10);
result = insertText(result, [10 10], num, 'BoxOpacity', 1, ...
    'FontSize', 14);
figure; imshow(result); title('***WARNING***SUSPICIOUS
ACTIVITY(RUNNING)');
videoPlayer = vision.VideoPlayer('Name', 'Detected Human');
videoPlayer.Position(3:4) = [700,500];  % window size: [width,
height]
se = strel('square', 5); % morphological filter for noise removal

while ~isDone(videoReader)

    frame = step(videoReader); % read the next video frame

    % Detect the foreground in the current video frame
    foreground = step(foregroundDetector, frame);

    % Use morphological opening to remove noise in the foreground
    filteredForeground = imopen(foreground, se);

    % Detect the connected components with the specified minimum
area, and
    % compute their bounding boxes
    bbox = step(blobAnalysis, filteredForeground);

    % Draw bounding boxes around the detected humans
    result = insertShape(frame, 'Rectangle', bbox, 'Color', 'red');

    % Display the number of humans found in the video frame
    num = size(bbox, 1);
    result = insertText(result, [10 10], num, 'BoxOpacity', 1, ...
```

```
        'FontSize', 14);

    step(videoPlayer, result);  % display the results
end

release(videoReader); % close the video file
```

Figure 2: object tracking (normal activity).

Figure 3: object tracking (suspicious activity).

CHAPTER 3

PROPOSED METHOD

In this chapter we will discuss the approach of our system that is examines the possibility of detecting suspicious activities and differentiating between them and those activities not suspicious, in our system the suspicious activity will be running, where we tried as much as possible to study all characteristics that related to this activity (running).

3.1 INPUT VIDEO:

This phase is first step in processing where the video input format is MOV, where the video consists of 23 frames per second, initial frame is reference frame in system. The reference frame is used to obtain the objects in image in next steps. Here, we did not dealt with any external datasets, where we used a set of videos as our dataset that is a set of videos that were captured by a fixed camera indoor scene.

3.2 EXTRACTING FRAMES:

As we know each video file consisted from a large number of frames N where the process of video processing and analysis is done on original raw frames, we also note that successive frames contain minor differences that are almost invisible only after the passage of a large number of frames. So video analysis is a complicated process that takes a very long time, especially if the video size is large as following in figure 4.

Figure 4: Extracting frames (convert video to frames).

3.3 BACKGROUND SUBTRCTION:

We used background subtraction method [22] for detect and separate the object (foreground regions) which is moving from the background frame (first frame), where we compare and subtract the current frame from reference frame. When we choose the reference frame we must attention to overlapping the image and try to avoid it, also the possibility of losing information from sequence of video that may great importance, background subtraction is represented as below:

$$|frameN-frameN-1| > Th \tag{3.1}$$

Where:

Frame N: the current frame.

Frame N-1: the reference frame.

Th: threshold value.

If the result of the subtraction is greater than Th value this means that the pixel is part of the foreground, otherwise it will be part of the background.

3.4 NOISE REMOVAL:

In last phase we applied the background subtraction method to get foreground objects but the results always contains a noise like shadows, reflections and lights....etc. So we must use a filter to reduce and eliminate this noise, for that we used Morphological operations in figure 5 where we apply opening (erosion followed by dilation) then closing(dilation followed by erosion) after that erosion(removes pixels on object boundaries) to get the better results.

Figure 5: Morphological operations.

3.5 FORGROUND DETECTION:

In this point after we got foreground objects also applied the morphological operations and get a clear image without any noise. In this phase we must do some mathematical operations to detection of various activities and the distinction between suspicious and non-suspicious activities.

3.5.1 Parallel with the camera:

The parallel motion of the object with the camera, whether described by walking or running is calculated by the absolute value of the x coordinate of centroid, if the displacement in x direction is small that means this motion is walking ,vice versa if it is large, that means the motion is running as follows:

$$X1 - X2 < T \text{ (for walking/ Normal).} \quad (3.2)$$

$$X1 - X2 > T \text{ (for running/ Suspicious).} \quad (3.3)$$

Where,

X1 is X- axis of previous frame.

X2 is X- axis of next frame.

T is threshold value.

3.5.2 To/away from the camera:

For this phase we must calculate change in the area of object between consecutive frames, based on change value we can know if the activity of this object is suspicious (big change/running) or normal activity (small change/walking), as below:

$$A = ((area1 - area2) / (area1 + area2)). \quad (3.4)$$

Where,

For (to the camera):

Area1 is size of the object in current frame.

Area2 is size of the object in previous frame.

As for (away from the camera):

Area1 is size of the object in current frame.

Area2 is size of the object in next frame.

We can see in figure 4 the results where first image is background image, second image is current frame that need to process, third, forth, and fifth images are results after applying the morphological operations and last image is the result of apply background subtraction method.

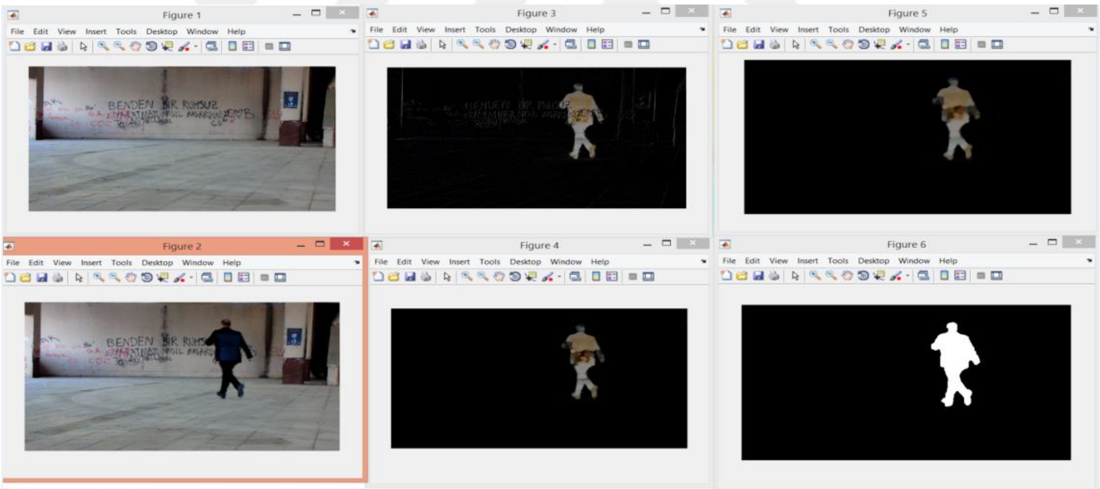


Figure 6: Background subtraction.

(figure1 (Background frame), figure2 (foreground frame), figure3, 4, 5(morphological operations), figure6 (the result "background subtraction")).

The following flowchart illustrates the structure of the proposed system.

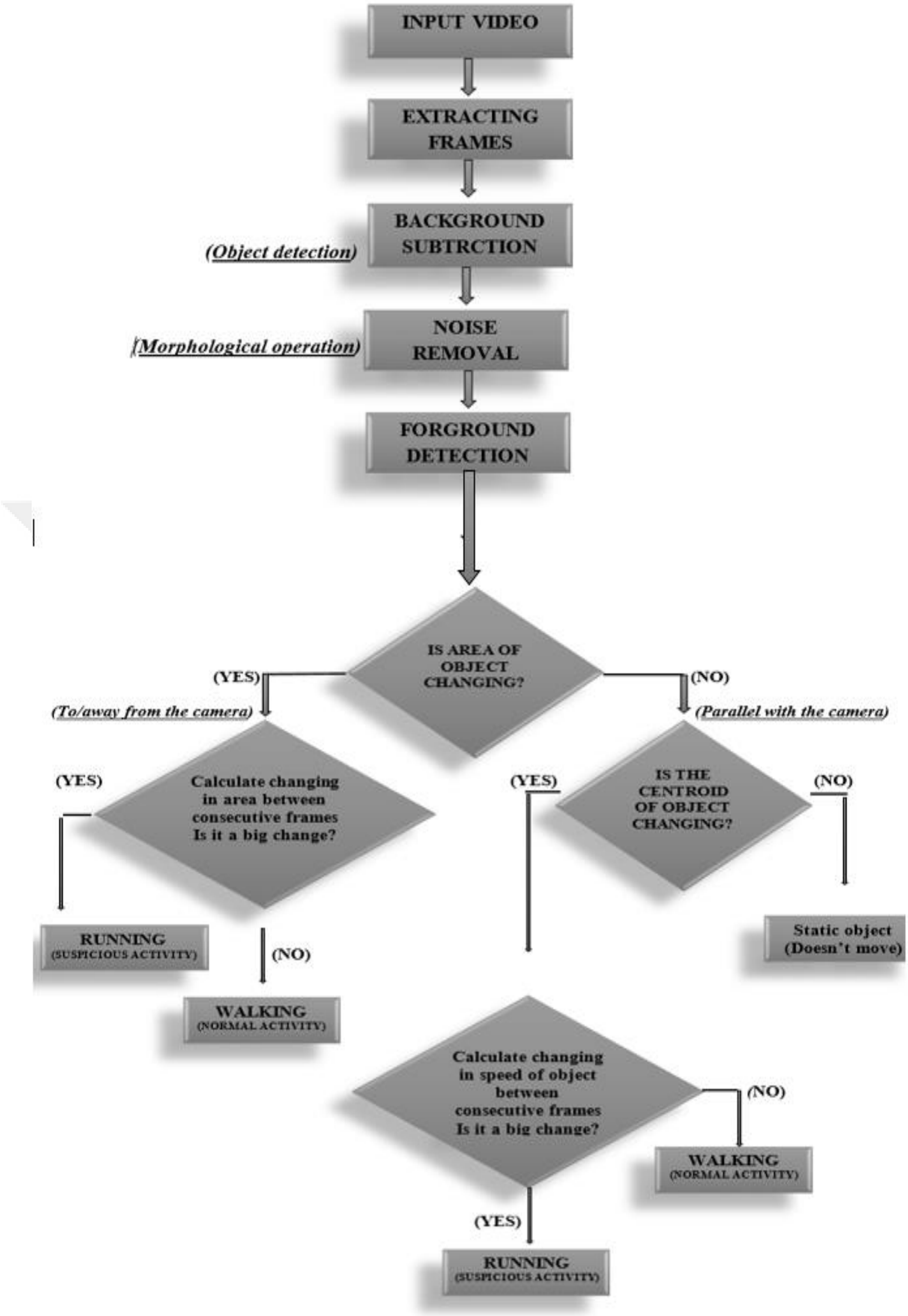


Figure 7: System Architecture.

CHAPTER 4

EXPERIMENTAL RESULTS

In this chapter we will discuss the experimental results obtained from the previous phases, which are walking and running.

In experimental set up we used a computer with core i7 processor, RAM with 2 GB and 64-bit operating system with windows 8.1. As for the programming environment was Matlab2016, we used our data set where we have captured some videos indoor, the videos that used were MOV format, the video consists of 23 frames per second.

4.1 Walking:

4.1.1 Parallel with the camera:

We have to find value of centroid to know if this object (human) is moving parallel with camera also to detect if it is moving from right to left or from left to right, we do this by calculated the absolute value of the x coordinate of centroid for consecutive frames, and compare the displacement value in x with threshold value to know if this activity is normal (walking) or suspicious (running), as below:

$$X1 - X2 < T \text{ (for walking/ Normal).} \quad (4.1)$$

$$X1 - X2 > T \text{ (for running/ Suspicious).} \quad (4.2)$$

Where,

X1 is x-axis of previous frame.

X2 is x-axis of next frame.

T is threshold value.

Example:

X-axis of (Frame#66-Frame#67 = 991.9535 -981.5305

= 10.4230).

4.1.1.1 Walking parallel with the camera (One object/Right to left/ near):

In table 1 we can observe the change in axis for successive frames, where the X-axis is decreases and Y-axes is almost fix, this indicates to that the object is moving from right to left, also the size of the object is big that's mean is the object is near from the camera, as well as we can conclude the motion of object in the frames and here the motion is walking because the change in the axis is considered a relative or simple change between successive frames.

Table 1: Results of walking one object (near) from right to left.

No of frame	**Area**	**Centre(x,y)**	**Frame rate**	**Walking**
66	61103	991.9535 365.5711	23 f/sec	10.4230
67	61239	981.5305 368.7759	23 f/sec	11.1177
68	61311	970.4128 382.6083	23 f/sec	13.7081
69	61597	957.7047 382.2494	23 f/sec	13.9633
70	61201	944.7414 381.9131	23 f/sec	12.3912
71	61364	932.3502 377.8981	23 f/sec	15.1293
72	61571	917.2209 378.3161	23 f/sec	13.4595
73	61120	901.7614 380.4759	23 f/sec	12.3555
74	61221	888.4059 380.0864	23 f/sec	16.7387
75	61311	876.6672 378.6266	23 f/sec	11.5511
76	61648	860.1161 378.5863	23 f/sec	11.4852
77	61451	849.6309 378.4627	23 f/sec	12.7100

78	61164	838.9209 370.9226	23 f/sec	10.9340
79	61215	826.9869 362.4377	23 f/sec	12.6155
80	61346	814.3714 384.6506	23 f/sec	#

(a) (b) (c)

Figure 8: Successive of frames

((a) Frame#66,(b)Frame#67, (c)Frame#68)).

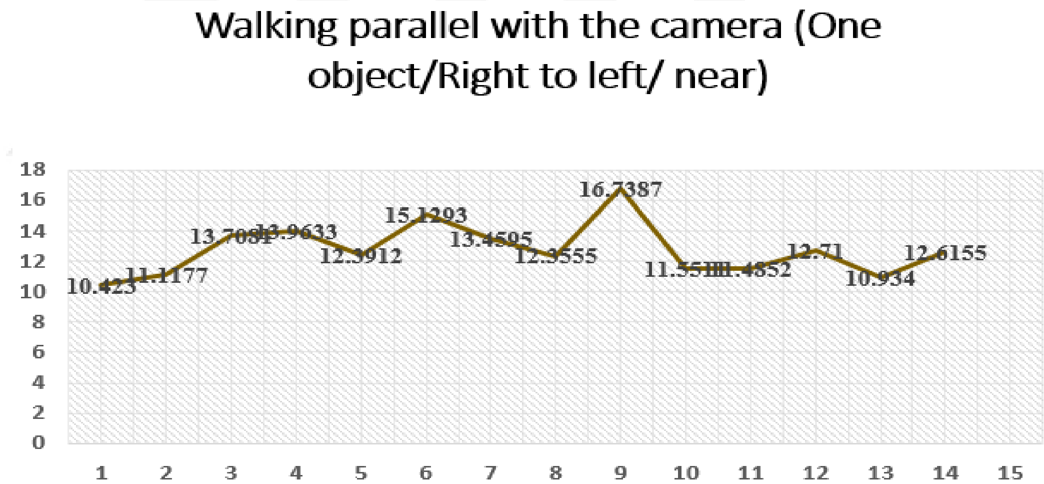


Figure 9: Curve of walking parallel with the camera (One object/Right to Left/ near).

4.1.1.2 Walking parallel with the camera (One object/Right to left/ far):

In table 2 we can observe the change in axes for successive frames, where the X-axis is decreases and Y-axis is almost fix, also the size of the object is small that’s mean is the object is far from the camera, this indicates to that the object is moving from right to left, as well as we can conclude the motion of object in the frames and here the motion is walking because the change in the axis is considered a relative or simple change between successive frames.

Table 2: Results of walking one object (far) from right to left.

No of frame	Area	Centre(x,y)	Frame rate	Walking
80	11694	1.0e+03 * 1.0363 0.3709	23 f/sec	10
81	11535	1.0e+03 * 1.0268 0.3697	23 f/sec	8
82	11680	1.0e+03 * 1.0188 0.3678	23 f/sec	9
83	11720	1.0e+03 * 1.0098 0.3680	23 f/sec	9
84	11876	1.0e+03 * 1.0009 0.3642	23 f/sec	9.7283
85	11670	990.2717 370.4688	23 f/sec	11.5826
86	11300	978.6891 375.4471	23 f/sec	9.8464
87	11383	968.8427 371.4718	23 f/sec	8.9974
88	11339	959.8453 372.1242	23 f/sec	7.8834
89	11514	951.9619 374.1164	23 f/sec	8.7452
90	11871	943.2167 371.2603	23 f/sec	8.9791
91	11654	934.2376 371.2694	23 f/sec	9.5855
92	11574	924.6521 370.9405	23 f/sec	10.8375
93	11344	913.8146 372.0364	23 f/sec	8.9402
94	11534	904.8744 371.2811	23 f/sec	7.2581
95	11461	897.6163 368.1307	23 f/sec	9.3011
96	11212	888.3152 364.9655	23 f/sec	9.1917
97	11562	879.1235 363.5669	23 f/sec	10.2826
98	11436	868.8409 372.6818	23 f/sec	10.7961
99	11337	858.0448 371.8266	23 f/sec	8.9809
100	11443	849.6390 370.5220	23 f/sec	#

(a) **(b)** **(c)**

Figure 10: Successive of frames

((a) Frame#85, (b) Frame#86, (c) Frame#87).

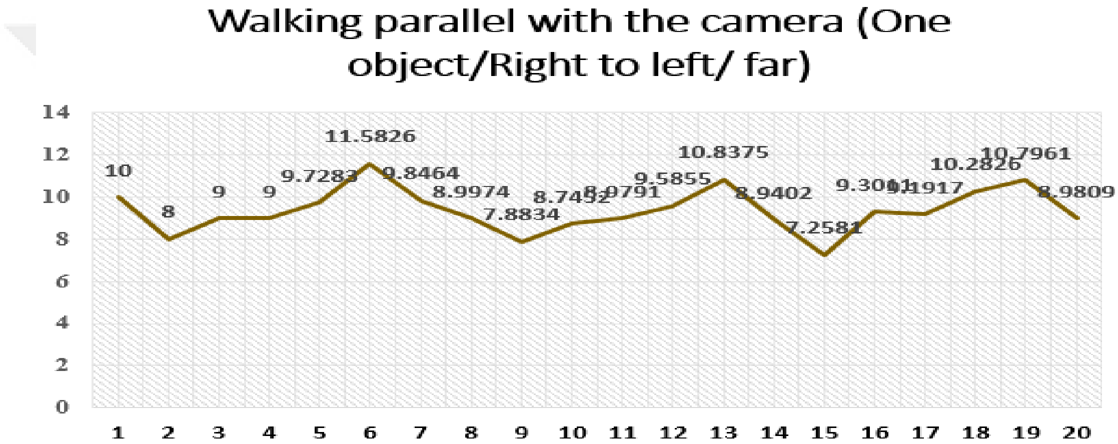


Figure 11: Curve of walking parallel with the camera (One object/Right to left/ far)

4.1.1.3 Walking parallel with the camera (One object/Left to right/ near):

In table 3 we can observe the change in axis for successive frames, where the X-axis is increases and Y-axis is almost fix, this indicates to that the object is moving from left to right, also the size of the object is big that’s mean is the object is near from the camera, as well as we can conclude the motion of object in the frames and here the motion is walking because the change in the axis is considered a relative or simple change between successive frames.

Table 3: Results of walking one object (near) from left to right.

No of frame	Area	Centre(x,y)	Frame rate	Walking
60	49726	223.2268 358.9311	23 f/sec	12.3995
61	49766	235.6233 356.6242	23 f/sec	11.3796
62	49676	246.0029 344.5216	23 f/sec	9.3071
63	49871	255.3100 352.7643	23 f/sec	12.8041
64	49699	267.1141 367.5435	23 f/sec	10.3916
65	49599	277.5057 363.7224	23 f/sec	9.7370
66	49425	286.2427 358.4278	23 f/sec	9.6818
67	49319	298.9245 352.1804	23 f/sec	11.2898
68	49496	307.2141 359.6451	23 f/sec	8.4173
69	49730	318.6314 351.4872	23 f/sec	12.7276
70	49381	326.3590 350.7493	23 f/sec	11.7125
71	49876	338.0715 352.0631	23 f/sec	12.6402
72	49653	350.7117 359.8894	23 f/sec	10.5492
73	49539	360.2609 358.9441	23 f/sec	11.5346
74	49727	371.7955 344.9700	23 f/sec	9.3185
75	49617	380.1140 369.2917	23 f/sec	11.7341
76	49397	391.8481 370.6394	23 f/sec	12.7292
77	49463	403.5773 373.0578	23 f/sec	9.2375
78	49315	412.8148 361.2782	23 f/sec	11.2784
79	49560	423.0932 363.6906	23 f/sec	#

(a) **(b)** **(c)**

Figure 12: Successive of frames

((a) Frame#60,(b) Frame#61, (c) Frame#62).

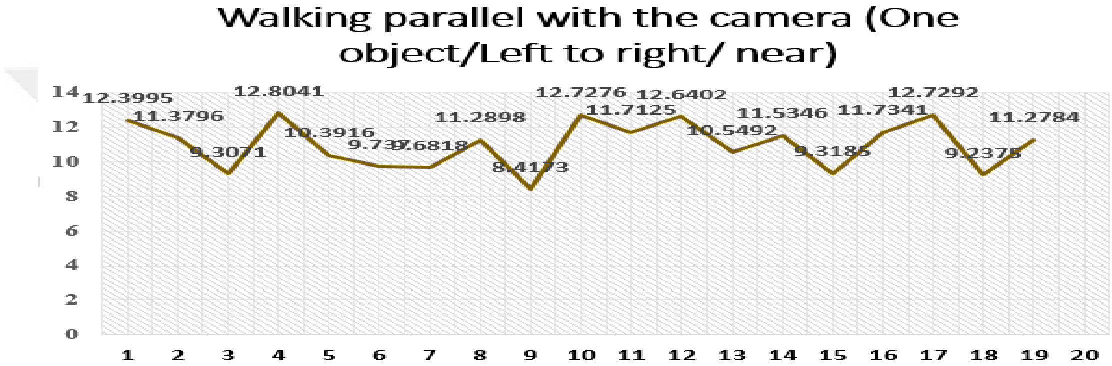


Figure 13: Curve of walking parallel with the camera (One object/Left to right/ near)

4.1.1.4 Walking parallel with the camera (One object/Left to right/ far):

In table 4 we can observe the change in axis for successive frames, where the X-axis is increases and Y-axis is almost fix, this indicates to that the object is moving from left to right, also the size of the object is small that's mean is the object is far from the camera, as well as we can conclude the motion of object in the frames and here the motion is walking because the change in the axis is considered a relative or simple change between successive frames.

Table 4: Results of walking one object (far) from left to right.

No of frame	Area	Centre(x,y)	Frame rate	Walking
45	23519	308.0265 363.4088	23 f/sec	9.703
46	22450	317.7295 361.0476	23 f/sec	11.6458

47	21256	329.3753 355.1640	23 f/sec	14.8052
48	22876	344.1805 355.9932	23 f/sec	12.833
49	25844	357.0135 365.6502	23 f/sec	12.8616
50	26062	369.8751 365.2297	23 f/sec	10.9994
51	26391	380.8745 364.4834	23 f/sec	9.7006
52	26677	390.5751 363.2271	23 f/sec	9.3792
53	27049	399.9543 362.5039	23 f/sec	10.8929
54	27612	410.8472 363.3547	23 f/sec	10.8228
55	27735	421.6700 363.7467	23 f/sec	11.451
56	27331	433.1210 363.6736	23 f/sec	12.0767
57	26131	445.1977 363.7301	23 f/sec	12.1450
58	24703	457.3427 362.2017	23 f/sec	11.1414
59	23776	468.4841 360.3218	23 f/sec	10.4998
60	22791	478.9839 357.1411	23 f/sec	11.2547
61	22696	490.2386 356.1349	23 f/sec	13.2996
62	25689	503.5382 368.7119	23 f/sec	13.8215
63	26168	517.3597 368.9369	23 f/sec	12.3748
64	25967	529.7345 367.4263	23 f/sec	11.0026
65	25477	540.7371 366.5493	23 f/sec	#

(a) **(b)** **(c)**

Figure 14: Successive of frames

((a) Frame#45,(b) Frame#46, (c) Frame#47).

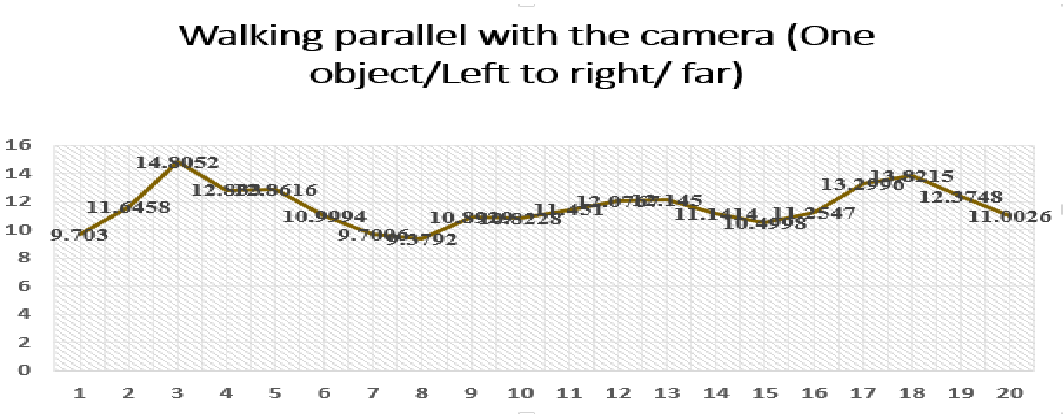


Figure 15: Curve of walking parallel with the camera (One object/Left to right/ far).

4.1.1.5 Walking parallel with the camera (Two objects/near):

In table 5 we can observe the change in axis for successive frames, where the X-axis of object(1)is increases and Y-axis of object(1) is almost fix, while the X-axis of object(2)is decreases and Y-axis of object(2) is also almost fix, this indicates to that object(1) is moving from left to right as for object(2) is moving from right to left, also the size of the objects are big that's mean is the objects are near from the camera, as well as we can conclude the motion of objects in the frames and here the motion is walking because the change in the axis is considered a relative or simple change between successive frames.

Table 5: Results of walking two objects (near).

No frame	Area Object (1)	Area Object (2)	Centre of object1(x,y)	Centre of object2(x,y)	Frame rate	Walking	
						Obj1	Obj2
59	65609	63120	226.742 339.819	1.0e+03 * 1.1653 0.3708	23f/sec	14.1218	20
60	66655	65155	240.864 347.207	1.0e+03 * 1.1451 0.3715	23f/sec	16.1778	16
61	67514	65398	257.042 354.102	1.0e+03 * 1.1298 0.3755	23f/sec	16.8272	22
62	62956	63456	273.869 347.160	1.0e+03 * 1.1079 0.3704	23f/sec	22.0095	17
63	60773	62895	295.878 346.394	1.0e+03 * 1.0904 0.3690	23f/sec	23.3321	18

64	57109	61238	319.210 324.827	1.0e+03 * 1.0720 0.3676	23f/sec	23.0884	20
65	65145	58509	342.299 344.276	1.0e+03 * 1.0526 0.3636	23f/sec	23.1489	18
66	70557	56378	365.448 357.513	1.0e+03 * 1.0341 0.3619	23f/sec	17.0084	20
67	71457	54271	382.456 353.609	1.0e+03 * 1.0143 0.3541	23f/sec	16.4365	20.4959
68	71917	51759	398.893 349.628	993.504 342.401	23f/sec	16.7736	22.6995
69	72034	55319	415.666 351.818	970.804 346.382	23f/sec	18.6593	22.6995
70	73369	63466	434.326 353.746	950.763 369.789	23f/sec	22.0975	20.0415
71	73607	64351	456.423 360.256	928.746 372.416	23f/sec	19.1922	22.0168
72	71154	64523	475.615 357.037	908.778 373.828	23f/sec	19.5896	19.9675
73	66940	64362	495.205 356.467	891.726 374.645	23f/sec	19.5653	17.0519
74	62901	64117	514.770 353.800	873.364 374.669	23f/sec	17.3279	18.3624
75	61559	62881	532.098 352.138	854.528 371.396	23f/sec	22.0949	18.8365
76	61609	61526	554.193 343.153	836.398 368.819	23f/sec	21.2883	18.1298
77	66244	57963	575.481 353.292	815.208 361.995	23f/sec	19.2061	21.1900
78	70834	55718	594.687 364.197	796.398 356.660	23f/sec	#	#

(a) **(b)** **(c)**

Figure 16: Successive of frames

((a) Frame#59,(b) Frame#60, (c) Frame#61).

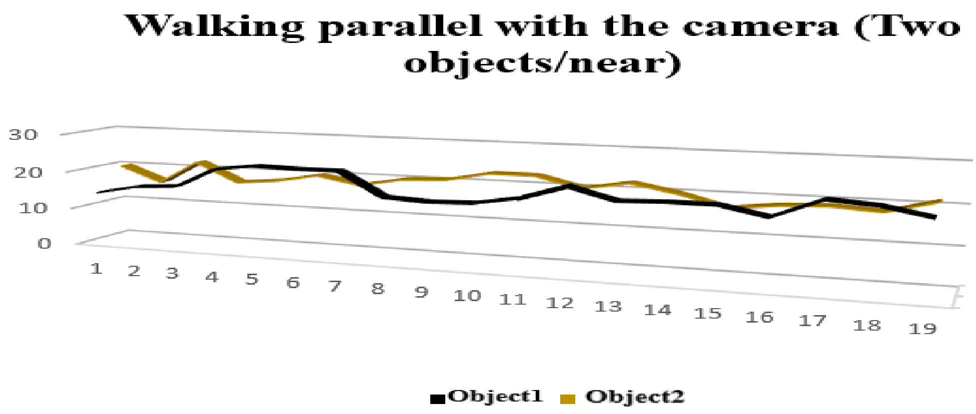


Figure 17: Curve of walking parallel with the camera (Two objects/near).

4.1.1.6 Walking parallel with the camera (Two objects/ middle):

In table 6 we can observe the change in axis for successive frames, where the X-axis of object(1)is increases and Y-axis of object(1) is almost fix ,while the X-axis of object(2)is decreases and Y-axis of object(2) is also almost fix ,this indicates to that object(1) is moving from left to right as for object(2) is moving from right to left, also the size of the objects are not big and not small that's mean is the objects are in the middle from the camera, as well as we can conclude the motion of objects in the frames and here the motion is walking because the change in the axis is considered a relative or simple change between successive frames.

Table 6: Results of walking two objects (middle).

No. frame	Area object(1)	Area object(2)	Centre of object1(x,y)	Centre of object2(x,y)	Frame rate	Walking Obj1	Walking Obj2
55	28053	27800	421.519 347.239	1.0e+03 * 1.2465 0.3530	23 f/sec	12.6126	14
56	28106	26279	434.132 345.891	1.0e+03 * 1.2326 0.3505	23 f/sec	11.9407	14
57	28272	24315	446.072 343.702	1.0e+03 * 1.2180 0.3409	23 f/sec	12.4649	16
58	28128	26542	458.537 340.128	1.0e+03 * 1.2021 0.3608	23 f/sec	13.1942	18
59	282099	28433	471.732 337.193	1.0e+03 * 1.1848 0.3701	23 f/sec	16.5000	16

60	28220	28230	488.232 326.703	1.0e+03 * 1.1686 0.3744	23 f/sec	14.3727	14
61	28335	29486	502.604 336.360	1.0e+03 * 1.1549 0.3631	23 f/sec	13.7151	11
62	28516	29501	516.319 344.666	1.0e+03 * 1.1431 0.3594	23 f/sec	13.0700	14
63	28659	29135	529.389 344.776	1.0e+03 * 1.1293 0.3612	23 f/sec	10.9050	12
64	28798	30560	540.294 344.403	1.0e+03 * 1.1173 0.3610	23 f/sec	11.2569	13
65	28687	30649	551.551 345.572	1.0e+03 * 1.1048 0.3591	23 f/sec	13.1054	12
66	28463	30255	564.657 345.539	1.0e+03 * 1.0925 0.3583	23 f/sec	13.4011	13
67	27533	28990	578.058 344.832	1.0e+03 * 1.0793 0.3549	23 f/sec	11.5991	14
68	28641	27994	589.657 343.120	1.0e+03 * 1.0655 0.3530	23 f/sec	11.7921	16
69	28515	26696	601.449 341.640	1.0e+03 * 1.0490 0.3430	23 f/sec	12.1565	18
70	28730	29069	613.605 340.564	1.0e+03 * 1.0316 0.3445	23 f/sec	14.1276	15
71	28476	32310	627.733 336.548	1.0e+03 * 1.0167 0.3573	23 f/sec	14.7073	15
72	28713	33055	642.440 336.186	1.0e+03 * 1.0014 0.3592	23 f/sec	13.3068	13.59
73	28829	32705	655.747 348.221	987.408 357.279	23 f/sec	14.6309	12.11
74	28942	32724	670.378 349.148	975.298 355.346	23 f/sec	11.8866	9.693
75	27132	32391	682.265 347.110	965.604 356.957	23 f/sec	#	#

(a) **(b)** **(c)**

Figure 18: Successive of frames

((a) Frame#55,(b) Frame#56, (c) Frame#57).

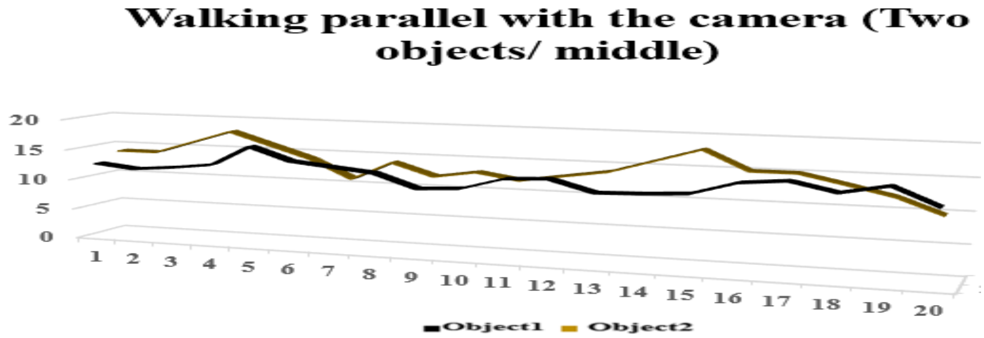


Figure 19: Curve of walking parallel with the camera (Two objects/ middle).

4.1.1.7 Walking parallel with the camera (Two objects/ far):

In table 7 we can observe the change in axis for successive frames, where the X-axis of object(1)is increases and Y-axis of object(1) is decreases, while the X-axis of object(2)is decreases and Y-axis of object(2) is increases ,this indicates to that object(1) is moving from left to right as for object(2) is moving from right to left, also the size of the objects is too small that's mean is the objects are in the far from the camera, as well as we can conclude the motion of objects in the frames and here the motion is walking because the change in the axis is considered a relative or simple change between successive frames.

Table 7: Results of walking two objects (far).

No. frame	Area of object (1)	Area of object (2)	Centre of object1(x,y)	Centre of object2(x,y)	Frame rate	Walking	
						Obj1	Obj2
45	14942	13959	384.187 352.388	1.0e+03 * 1.0287 0.3611	23f/sec	9.8111	9
46	14908	13884	393.998 351.766	1.0e+03 * 1.0194 0.3594	23f/sec	8.6894	11
47	14789	13768	402.687 352.729	1.0e+03 * 1.0083 0.3527	23f/sec	9.2097	9.6502
48	14748	13890	411.897 353.553	998.349 358.795	23f/sec	9.7194	9.2540
49	14904	13796	421.617 352.776	989.095 366.478	23f/sec	8.8564	9.4295
50	14752	13608	430.473 353.244	979.666 365.292	23f/sec	9.9014	7.373
51	14841	13776	440.374 351.128	972.293 364.360	23f/sec	8.3863	6.9687
52	14750	13896	448.761 350.131	965.324 362.615	23f/sec	10.0429	6.2372
53	14939	13747	458.804 347.376	959.087 362.977	23f/sec	10.9311	7.5696
54	14791	13675	469.735 344.109	951.517 362.964	23f/sec	10.6626	8.5541
55	14888	13528	480.397 347.725	942.963 363.708	23f/sec	10.6067	8.8488
56	14737	13784	491.004 358.671	934.114 363.433	23f/sec	10.4471	8.2169
57	14611	13895	501.451 356.202	925.898 363.252	23f/sec	9.663	8.1425
58	14425	13785	511.114 355.937	917.755 360.938	23f/sec	9.737	7.4882
59	14584	13955	520.851 357.054	910.267 360.817	23f/sec	9.0623	10.941
60	14798	13775	529.913 357.508	899.326 353.267	23f/sec	8.9294	10.572
61	14562	13816	538.843 358.086	888.754 354.934	23f/sec	9.0995	9.8998
62	14340	13929	547.942 358.233	878.854 365.714	23f/sec	9.0569	9.5187
63	13493	14042	556.999 357.050	869.335 363.739	23f/sec	9.9654	7.7915
64	13360	14137	566.965 353.479	861.544 363.794	23f/sec	#	#

(a) **(b)** **(c)**

Figure 20: Successive of frames

((a) Frame#45,(b) Frame#46, (c) Frame#47).

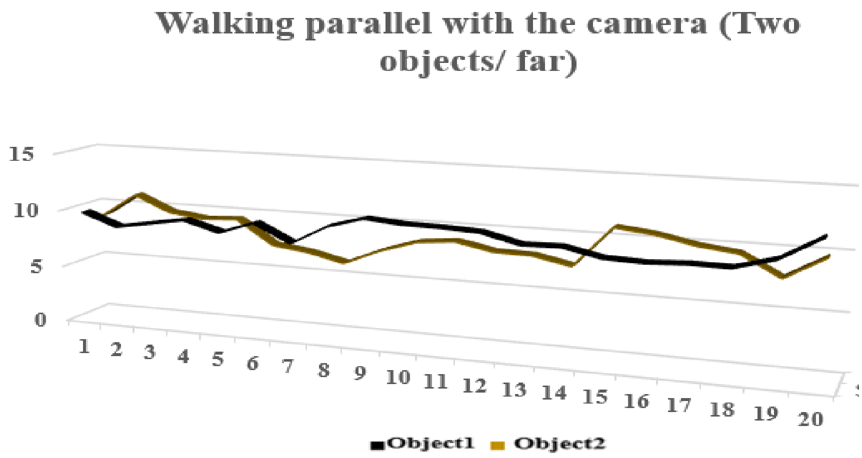


Figure 21: Curve of walking parallel with the camera (Two objects/ far).

4.1.2 To/ Away from the camera:

In phase of (To/away) from the camera, we computed the change area for object size between consecutive frames by use:

$$A= (|\ (area1 - area2)\ |/\ (area1+area2)). \quad (4.3)$$

Where,

A <T (for walking/ Normal).

A >T (for running/ Suspicious).

T is threshold value.

Example:

T= Default value.

Size of ((Frame#22-Frame#21)/ (Frame#22+Frame#21))

((20522-20470)/ (20522+20470))

= 0.0013.

If A < T that mean the activity is normal (walking).

A > T that mean the activity is suspicious (running).

4.1.2.1 Walking to the camera (One object):

In table 8 we can observe the change in size between successive frames, where the size of object increases when it approaches the camera and this indicates to that the object is moving towards the camera, as well as we can conclude the motion in the frames and here the motion is walking because the increase in size of object is considered a relative or simple increase between successive frames.

Table 8: Results of walking one object to the camera.

No of frame	Area	Centre(x,y)	Frame rate	Walking
21	20470	683.1158 363.0640	23 f/sec	0.0013
22	20522	683.8530 362.3088	23 f/sec	0.0035
23	20668	684.3574 361.1804	23 f/sec	0.0031
24	20798	684.2397 359.5966	23 f/sec	0.0034
25	20941	684.4482 359.1196	23 f/sec	0.0033
26	21080	683.2540 358.5813	23 f/sec	0.0048
27	21283	683.0613 358.3855	23 f/sec	0.0028
28	21401	682.0192 359.4144	23 f/sec	0.0032
29	21539	680.7539 360.9978	23 f/sec	0.0040
30	21713	680.4723 363.9863	23 f/sec	0.0038
31	21879	679.5840 366.1875	23 f/sec	0.0007
32	21911	678.2155 364.3147	23 f/sec	0.0041
33	22092	677.5545 365.1696	23 f/sec	0.0030
34	22224	676.2554 366.9817	23 f/sec	0.0031
35	22364	675.2836 367.8302	23 f/sec	0.0044

36	22562	673.5037 366.1471	23 f/sec	0.0009
37	22606	672.9512 362.6579	23 f/sec	0.0049
38	22828	671.6052 359.7243	23 f/sec	0.0019
39	22914	670.4744 358.8047	23 f/sec	0.0057
40	23179	670.8444 358.8000	23 f/sec	#

(a) **(b)** **(c)**

Figure 22: Example of successive frames ((a)Frame#21, (b) Frame#22, (c) Frame#23).

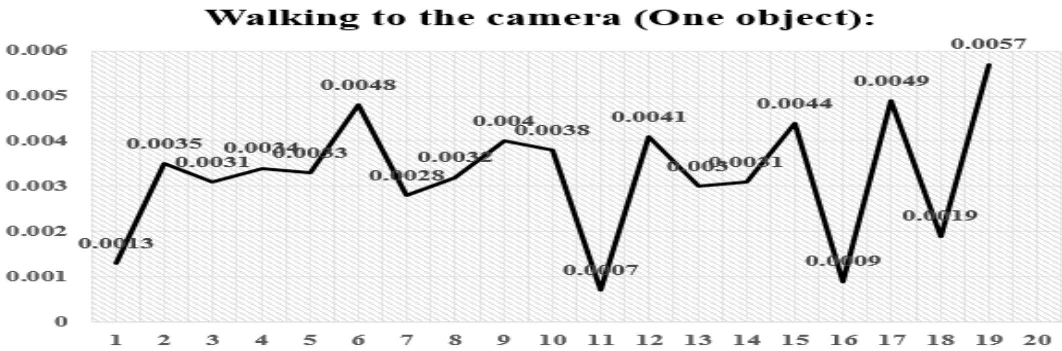


Figure 23: Curve of walking to the camera (One object).

4.1.2.2 Walking to the camera (Two objects):

In table 9 we can observe the change in size between successive frames, where the size of objects are increases when they approaches the camera and this indicates to that the objects are moving towards the camera, as well as we can conclude the motion in the frames and here the motion is walking because the increase in size of objects is considered a relative or simple increase between successive frames.

Table 9: Results of walking two objects to the camera.

No. frame	Area of object(1)	Area of object(2)	Centre of object1(x,y)	Centre of object2(x,y)	Frame rate	Walking	
						Obj1	Obj2
120	26785	29135	656.855 307.177	1.0e+03 * 1.0233 0.3327	23f/sec	0.0013	0.0040
121	26856	29370	655.218 307.008	1.0e+03 * 1.0232 0.3339	23f/sec	0.0059	0.0024
122	27172	29512	653.763 306.457	1.0e+03 * 1.0232 0.3338	23f/sec	0.0023	0.0033
123	27297	29707	651.794 308.213	1.0e+03 * 1.0232 0.3350	23f/sec	0.0018	0.0075
124	27393	30153	650.494 308.118	1.0e+03 * 1.0229 0.3375	23f/sec	0.0031	0.0012
125	27566	30224	649.052 307.263	1.0e+03 * 1.0210 0.3360	23f/sec	0.0014	0.0010
126	27643	30286	647.434 306.612	1.0e+03 * 1.0209 0.3391	23f/sec	0.0021	0.0004
127	27762	30311	645.641 305.319	1.0e+03 * 1.0217 0.3380	23f/sec	0.0008	0.0023
128	27807	30453	644.530 308.874	1.0e+03 * 1.0226 0.3384	23f/sec	0.0149	0.0019
129	28648	30568	643.708 310.097	1.0e+03 * 1.0213 0.3373	23f/sec	0.0028	0.0010
130	28809	30631	639.051 310.990	1.0e+03 * 1.0239 0.3322	23f/sec	0.0029	0.0023
131	28978	30772	637.960 312.582	1.0e+03 * 1.0249 0.3312	23f/sec	0.0025	0.0008
132	29126	30823	636.066 315.701	1.0e+03 * 1.0246 0.3320	23f/sec	0.0022	0.0019
133	29254	30938	636.119 310.740	1.0e+03 * 1.0248 0.3298	23f/sec	0.0029	0.0022
134	29426	31072	633.405 313.123	1.0e+03 * 1.0256 0.3336	23f/sec	0.0024	0.0018
135	29569	31181	629.446 316.547	1.0e+03 * 1.0260 0.3310	23f/sec	0.0021	0.0009
136	29692	31242	628.415 316.080	1.0e+03 * 1.0271 0.3328	23f/sec	0.0023	0.0022
137	29828	31378	623.005 316.247	1.0e+03 *	23f/sec	0.0022	0.0005

				1.0284 0.3344			
138	29959	31412	621.745 309.738	1.0e+03 * 1.0280 0.3353	23f/sec	0.0033	0.0023
139	30159	31559	617.018 308.339	1.0e+03 * 1.0287 0.3370	23f/sec	0.0020	0.0012

(a) **(b)** **(c)**

Figure 24: Example of successive frames

((a)Frame#120, (b) Frame#121, (c) Frame#122).

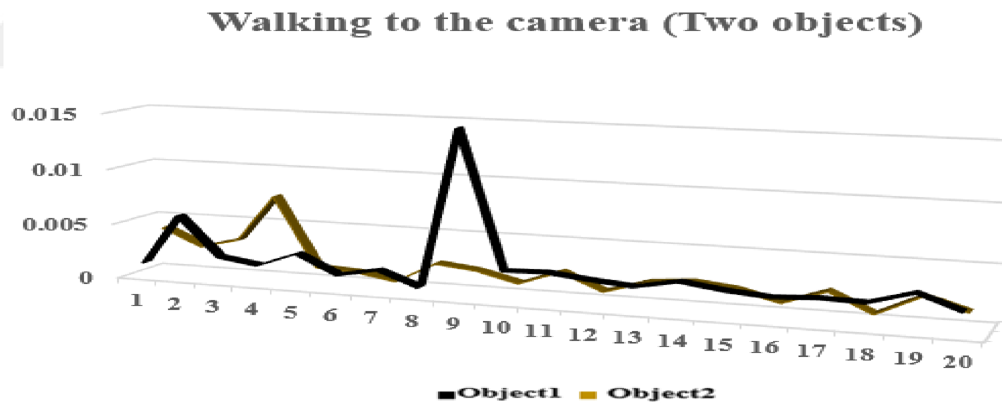


Figure 25: Curve of walking to the camera (Two objects).

4.1.2.3 Walking away from the camera (Two objects) :

In table 10 we can observe the change in size between successive frames, where the size of objects are decreases, this indicates to that the objects are moving away from the camera, as well as we can conclude the motion in the frames and here the motion is walking because the increase in size of objects are considered a relative or simple increase between successive frames.

Table 10: Results of walking two objects away from the camera.

No. frame	**Area of obj(1)**	**Area of obj(2)**	**Centre of object1(x,y)**	**Centre of object2(x,y)**	**Frame rate**	**Walking**	
						Obj1	**Obj2**
101	74582	66770	298.727 310.054	1.0e+03 * 1.0443 0.3549	23 f/sec	0.0006	0.0012
102	74486	66607	308.392 318.180	1.0e+03 * 1.0434 0.3641	23 f/sec	0.0011	0.0001
103	74324	66586	317.891 322.187	1.0e+03 * 1.0419 0.3705	23 f/sec	0.0007	0.0013
104	74220	66410	329.525 327.386	1.0e+03 * 1.0356 0.3746	23 f/sec	0.0010	0.0009
105	74070	66287	339.678 328.299	1.0e+03 * 1.0305 0.3748	23 f/sec	0.0011	0.0008
106	73909	66180	350.160 326.283	1.0e+03 * 1.0234 0.3725	23 f/sec	0.0008	0.0012
107	73777	66019	361.253 324.475	1.0e+03 * 1.0181 0.3677	23 f/sec	0.0006	0.0002
108	73633	65986	372.637 322.094	1.0e+03 * 1.0131 0.3609	23 f/sec	0.0009	0.0002
109	73537	65848	385.406 319.085	1.0e+03 * 1.0107 0.3558	23 f/sec	0.0006	0.0010
110	73426	65751	393.382 327.553	1.0e+03 * 1.0076 0.3509	23 f/sec	0.0007	0.0007
111	73330	65610	413.751 338.969	1.0e+03 * 1.0030 0.3542	23 f/sec	0.0006	0.0011
112	73124	65500	420.459 340.395	1.0e+03 * 1.0053 0.3635	23 f/sec	0.0014	0.0008
113	73008	65435	427.502 347.349	1.0e+03 * 1.0004 0.3634	23 f/sec	0.0007	0.0004

114	72919	65357	437.379 337.439	999.843 367.900	23 f/sec	0.0006	0.0005
115	72796	65199	443.572 344.170	990.732 364.237	23 f/sec	0.0008	0.0012
116	72644	65026	453.366 346.540	987.514 371.925	23 f/sec	0.0010	0.0013
117	72488	64931	465.797 357.482	983.471 368.133	23 f/sec	0.0011	0.0007
118	72320	64801	467.308 353.457	985.738 371.048	23 f/sec	0.0012	0.0010
119	72264	64658	477.666 355.093	986.221 365.128	23 f/sec	0.0003	0.0011
120	72127	64457	486.694 349.292	988.351 370.983	23 f/sec	#	#

(a) **(b)** **(c)**

Figure 26: Example of successive frames.

((a)Frame#101, (b) Frame#102, (c) Frame#103).

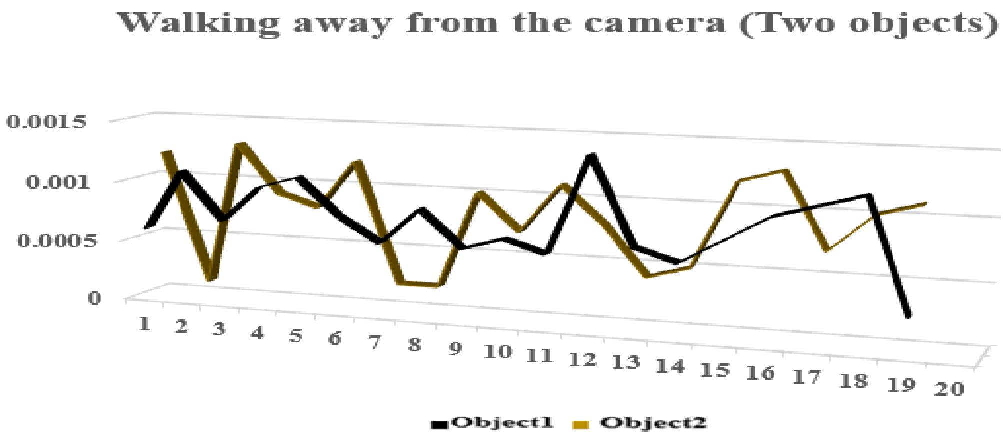


Figure 27: Curve of walking away from the camera (Two objects).

4.2 Running:

In running case we can note the big change in x-axes between successive frames, that big change indicate to the object (human) in running case, the value of x-axis for (frame#47) was 720.0596 for first object and 682.7851 for (frame#48) , this huge change in the X-axis proves that activity is a suspicious activity (running).

We also use the same Mathematical Equation that is:

- For (parallel cases):

$$X1 - X2 < T \text{ (for walking/ Normal).} \tag{4.5}$$

$$X1 - X2 > T \text{ (for running/ Suspicious).} \tag{4.6}$$

(Where “T” is a default value).

4.2.1 Parallel with the camera:

4.2.1.1 Running parallel with the camera (Two objects /near) :

In table 11 we can observe the change in axis for successive frames, where the X-axis of object(1) is increases and Y-axis of object(1) is almost fix, while the X-axis of object(2)is decreases and Y-axis of object(2) is also almost fix ,this indicates to that object(1) is moving from left to right as for object(2) is moving from right to left, also the size of the objects are too big that’s mean is the objects are near from the camera, as well as we can conclude the motion in the frames and here the motion is running because the change in the axis is a big change between successive frames.

Table 11: Results of running two near objects parallel with the camera.

No. frame	Area of obj(1)	Area of obj(2)	Centre of object1(x,y)	Centre of object2(x,y)	Frame rate	Running	
						Obj1	Obj2
47	67515	53735	720.059 336.986	923.759 374.686	23 f/sec	37.2745	33.5484

48	67796	53865	682.785 348.642	957.307 368.099	23 f/sec	42.0888	33.175
49	67774	54077	640.696 337.660	990.482 377.452	23 f/sec	39.8001	37.5175
50	67857	53915	600.896 326.709	1.0e+03 * 1.0281 0.3715	23 f/sec	39.5764	34
51	67757	53761	561.319 322.027	1.0e+03 * 1.0623 0.3689	23 f/sec	42.2485	31
52	67677	53886	519.071 327.609	1.0e+03 * 1.0935 0.3665	23 f/sec	41.4866	35
53	67549	54200	477.584 336.174	1.0e+03 * 1.1284 0.3710	23 f/sec	41.4117	32
54	67199	54164	436.167 338.872	1.0e+03 * 1.1619 0.3767	23f/sec	38.2619	33
55	67397	54321	397.905 341.804	1.0e+03 * 1.1947 0.3825	23 f/sec	46.3397	34
56	67599	54636	351.566 336.686	1.0e+03 * 1.2280 0.3867	23 f/sec	37.3762	37
57	67668	54342	314.189 349.808	1.0e+03 * 1.2654 0.3835	23 f/sec	#	#

(a) **(b)** **(c)**

Figure 28: Example of successive frames

((a)Frame#46, (b) Frame#47, (c) Frame#48).

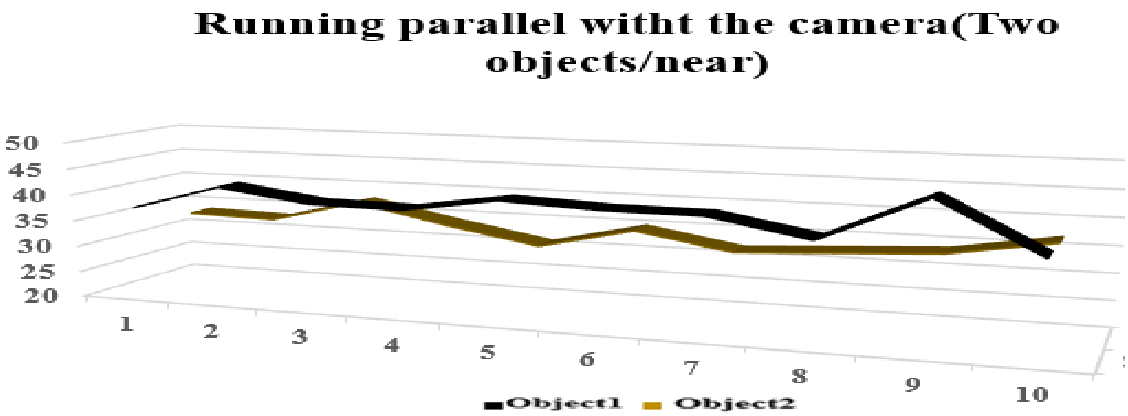


Figure 29: Curve of running parallel with the camera (Two objects /near).

4.2.1.2 Running parallel with the camera (Two objects /Middle) :

In table 12 we can observe the change in axis for successive frames, where the X-axis of object(1)is increases and Y-axis of object(1) is almost fix, while the X-axis of object(2)is decreases and Y-axis of object(2) is also almost fix ,this indicates to that object(1) is moving from left to right as for object(2) is moving from right to left, also the size of the objects are not big and not small that's mean is the objects are in the middle from the camera, as well as we can conclude the motion in the frames and here the motion is running because the change in the axis is a big change between successive frames.

Table 12: Results of running two objects (middle) parallel with the camera.

No. frame	Area of obj(1)	Area of obj(2)	Centre of object1(x,y)	Centre of object2(x,y)	Frame rate	Running Obj1	Running Obj2
60	28147	25415	306.648 367.027	1.0e+03 * 1.2787 0.3432	23 f/sec	42.718	39
61	28337	25516	349.366 367.258	1.0e+03 * 1.2397 0.3493	23 f/sec	42.656	41
62	28442	25600	392.023 363.510	1.0e+03 * 1.1985 0.3567	23 f/sec	41.116	46
63	28155	25486	433.139 364.579	1.0e+03 * 1.1528 0.3564	23 f/sec	43.043	24

64	28479	25219	476.183 372.759	1.0e+03 * 1.1284 0.3488	23 f/sec	41.823	63
65	28620	25599	518.006 374.739	1.0e+03 * 1.0653 0.3469	23 f/sec	28.519	34
66	28586	25400	546.526 370.752	1.0e+03 * 1.0312 0.3516	23 f/sec	44.322	31
67	28548	25309	590.848 370.284	1.0e+03 * 1.0001 0.3571	23 f/sec	42.941	47.7496
68	28763	25517	633.790 367.415	952.250 335.380	23 f/sec	36.031	32.8416
69	28673	25318	669.821 368.071	919.408 348.392	23 f/sec	38.184	38.3836
70	28764	25556	708.006 374.739	881.025 356.352	23 f/sec	39.519	34.9599
71	28893	25499	747.526 370.752	846.065 349.346	23 f/sec	36.319	38.6985
72	28955	25228	783.845 375.517	807.366 351.517	23 f/sec	39.003	41.6059
73	29023	25403	822.848 370.284	765.760 357.252	23 f/sec	32.973	37.3935
74	28971	25584	855.821 368.071	728.367 349.253	23 f/sec	37.968	36.0202
75	28794	25775	893.790 367.415	692.347 350.543	23 f/sec	#	#

(a) **(b)** **(c)**

Figure 30: Example of successive frames

((a)Frame#60, (b) Frame#61, (c) Frame#62).

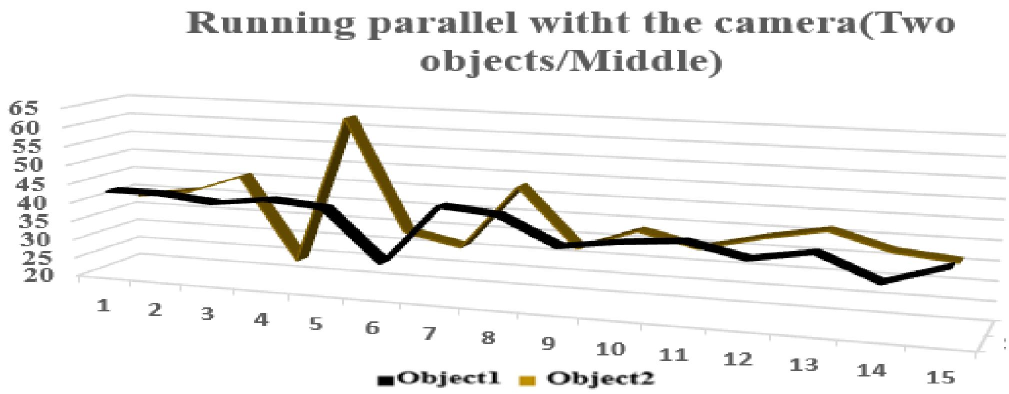


Figure 31: Curve of running parallel with the camera (Two objects /Middle).

4.2.1.3 Running parallel with the camera (Two objects /Far) :

In table 13 we can observe the change in axis for successive frames, where the X-axis of object(1)is increases and Y-axis of object(1) is almost fix, while the X-axis of object(2)is decreases and Y-axis of object(2) is also almost fix ,this indicates to that object(1) is moving from left to right as for object(2) is moving from right to left, also the size of the objects are small that's mean is the objects are far from the camera, as well as we can conclude the motion in the frames and here the motion is running because the change in the axis is a big change between successive frames.

Table 13: Results of running two objects (far) parallel with the camera.

No. frame	Area of obj(1)	Area of obj(2)	Centre of object1(x,y)	Centre of object2(x,y)	Fram e rate	Running	
						Obj1	Obj2
70	12267	11341	604.002 370.767	1.0e+03 * 1.1462 0.3601	23 f/sec	43.9497	45
71	12393	11267	647.951 364.657	1.0e+03 * 1.1011 0.3613	23 f/sec	42.8897	44
72	12359	11343	690.841 367.087	1.0e+03 * 1.0576 0.3565	23 f/sec	38.2522	36
73	12428	11441	729.093 368.964	1.0e+03 * 1.0218 0.3590	23 f/sec	39.3690	36.143
74	12294	11530	768.462 371.538	984.856 354.951	23 f/sec	43.3567	34.858

75	12316	11375	811.819 368.280	949.998 358.322	23 f/sec	35.1996	35.051
76	12376	11161	847.019 365.122	914.947 353.453	23 f/sec	39.1066	36.663
77	12337	11347	886.125 363.245	878.283 355.477	23 f/sec	40.2797	38.724
78	12277	11259	926.405 365.359	839.559 358.841	23 f/sec	37.0024	39.971
79	12462	11430	963.407 368.945	799.587 357.686	23 f/sec	37.5922	41.069
80	12259	11398	1.0e+03 * 1.0019 0.3677	758.518 358.102	23 f/sec	39	35.618
81	12383	11192	1.0e+03 * 1.040 366.914	722.899 358.467	23 f/sec	35	40.432
82	12495	11233	1.0e+03 * 1.075 367.433	682.467 355.582	23 f/sec	38	41.414
83	12239	11501	1.0e+03 * 1.113 370.293	641.052 356.447	23 f/sec	41	38.585
84	12376	11476	1.0e+03 * 1.154 370.143	602.467 354.582	23 f/sec	#	#

(a) **(b)** **(c)**

Figure 32: Example of successive frames

((a)Frame#70, (b) Frame#71, (c) Frame#72).

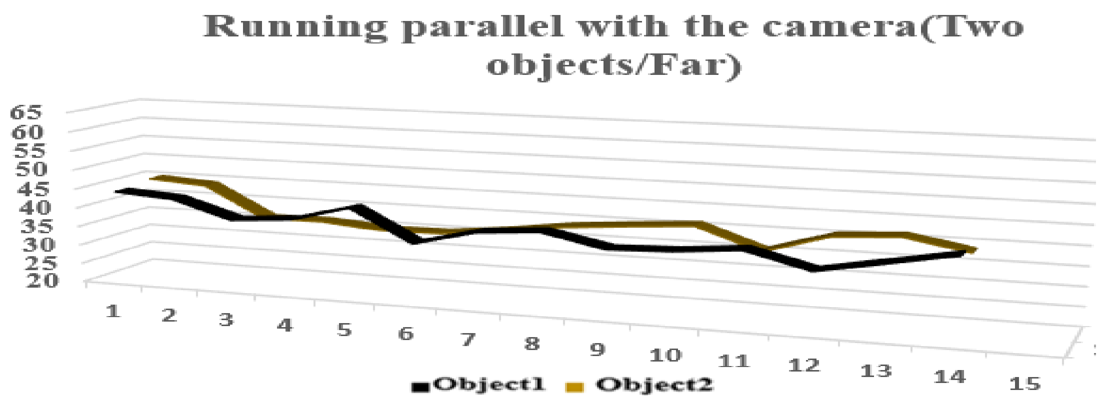


Figure 33: Curve of running parallel with the camera (Two objects /Far).

4.2.2 To/ Away from the camera:

4.2.2.1 Running away from the camera (Two objects):

In table 14 we can observe the change in size between successive frames, where the size of object is decreases, this indicates to that the object is a moving away from the camera, as well as we can conclude the motion of objects in the frames and here the motion is running because the change in the size of objects is a big change between successive frames.

Table 14: Results of running two objects away from the camera.

No. frame	Area of object(1)	Area of object(2)	Centre of object1(x,y)	Centre of object2(x,y)	Frame rate	Walking	
						Obj1	Obj2
46	55456	54818	283.633 361.536	1.0e+03 * 1.1773 0.3790	23 f/sec	0.0121	0.0145
47	54127	53247	311.887 362.968	1.0e+03 * 1.1628 0.3734	23 f/sec	0.0051	0.0090
48	53576	52294	335.118 355.330	1.0e+03 * 1.1500 0.3743	23 f/sec	0.0099	0.0102

49	52526	51241	362.345 350.386	1.0e+03 * 1.1421 0.3757	23 f/sec	0.0104	0.0106
50	51448	50170	386.471 353.493	1.0e+03 * 1.1302 0.3655	23 f/sec	0.0104	0.0112
51	50390	49062	409.684 358.096	1.0e+03 * 1.1236 0.3612	23 f/sec	0.0107	0.0111
52	49319	47981	433.518 354.992	1.0e+03 * 1.1131 0.3520	23 f/sec	0.0109	0.0107
53	48255	46965	455.300 355.267	1.0e+03 * 1.1033 0.3482	23 f/sec	0.0109	0.0113
54	47211	45914	478.284 356.990	1.0e+03 * 1.0943 0.3431	23 f/sec	0.0117	0.0112
55	46119	44895	491.684 357.575	1.0e+03 * 1.0858 0.3431	23 f/sec	0.0116	0.0122
56	45061	43815	503.459 351.475	1.0e+03 * 1.0763 0.3422	23 f/sec	0.0119	0.0124
57	43999	42741	518.398 326.571	1.0e+03 * 1.0638 0.3556	23 f/sec	0.0124	0.0120
58	42921	41731	538.000 338.575	1.0e+03 * 1.0584 0.3673	23 f/sec	0.0131	0.0128
59	41811	40678	553.706 336.871	1.0e+03 * 1.0490 0.3634	23 f/sec	0.0132	0.0128
60	40723	39646	567.087 332.748	1.0e+03 * 1.0442 0.3635	23 f/sec	#	#

(a) **(b)** **(c)**

Figure 34: Example of successive frames

((a)Frame#46, (b) Frame#47, (c) Frame#48).

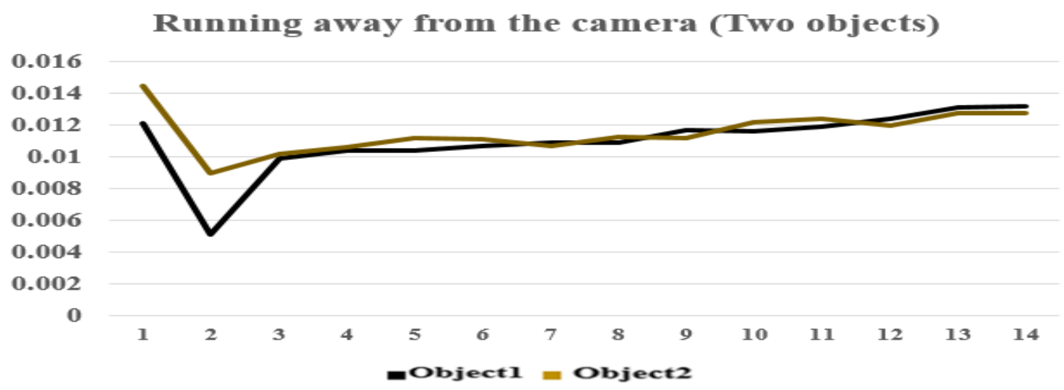


Figure 35: Curve of running away from the camera (Two objects).

4.2.2.2 Running to the camera(Two objects):

In table 15 we can observe the change in size between successive frames, where the size of objects are increases when they approaches the camera and this indicates to that the objects are moving towards the camera, as well as we can conclude the motion in the frames and here the motion is running because the increase in size of objects is a big increase between successive frames.

Table 15: Results of running of two objects to the camera.

No. frame	Area of object(1)	Area of object(2)	Centre of object1(x,y)	Centre of object2(x,y)	Frame rate	Running	
						Obj1	Obj2
40	11048	11003	786.471 334.169	967.695 343.485	23 f/sec	0.0468	0.0440
41	12134	12016	788.700 338.679	967.253 343.936	23 f/sec	0.0398	0.0406
42	13140	13032	788.167 338.769	966.663 341.310	23 f/sec	0.0387	0.0378
43	14198	14057	789.411 338.461	960.915 349.696	23 f/sec	0.0333	0.0363
44	15176	15115	782.810 334.055	966.439 353.014	23 f/sec	0.0304	0.0336
45	16128	16167	786.310 338.588	968.438 360.993	23 f/sec	0.0276	0.0319
46	17043	17232	783.836 347.800	965.871 367.231	23 f/sec	0.0328	0.0305
47	18198	18316	783.477 354.018	964.288 365.318	23 f/sec	0.0250	0.0243

48	19133	19229	783.165 354.903	968.916 368.567	23 f/sec	0.0260	0.0270
49	20154	20295	780.456 360.250	967.772 369.877	23 f/sec	0.0241	0.0274
50	21148	21440	779.633 373.962	964.286 377.150	23 f/sec	0.0260	0.0229
51	22276	22446	781.869 360.033	959.746 378.581	23 f/sec	0.0228	0.0241
52	23316	23554	769.118 369.283	975.850 366.906	23 f/sec	0.0211	0.0142
53	24319	24232	770.787 359.289	963.578 368.217	23 f/sec	0.0210	0.0206
54	25361	25252	774.374 354.294	967.938 369.300	23 f/sec	#	#

(a) (b) (c)

Figure 36: Example of successive frames

((a)Frame#40, (b) Frame#41, (c) Frame#42).

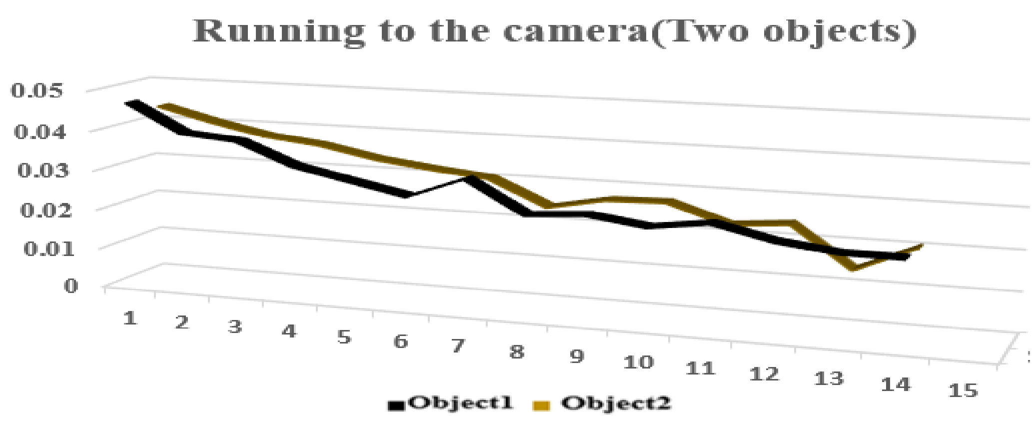


Figure 37: Curve of running to the camera (Two objects).

In next figure (38) reflected to us the GUI window for this system:

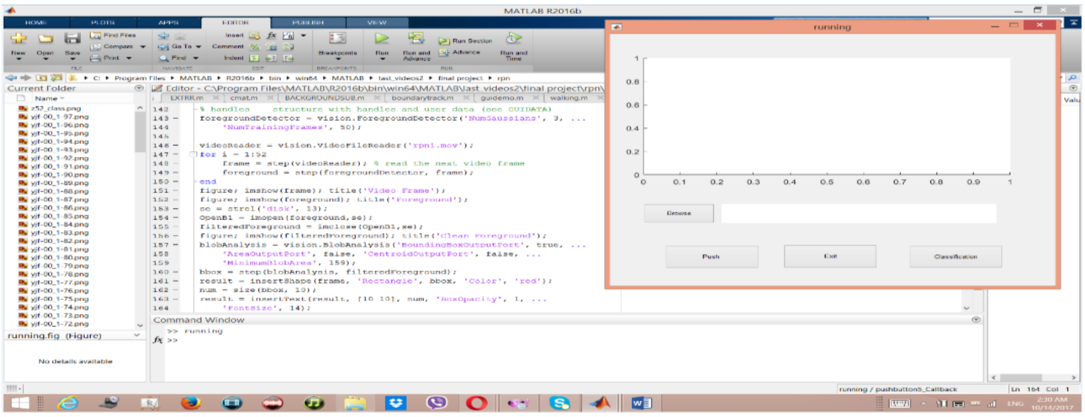


Figure 38: GUI window.

After applied this system the GUI window will be as shown in figure(39).

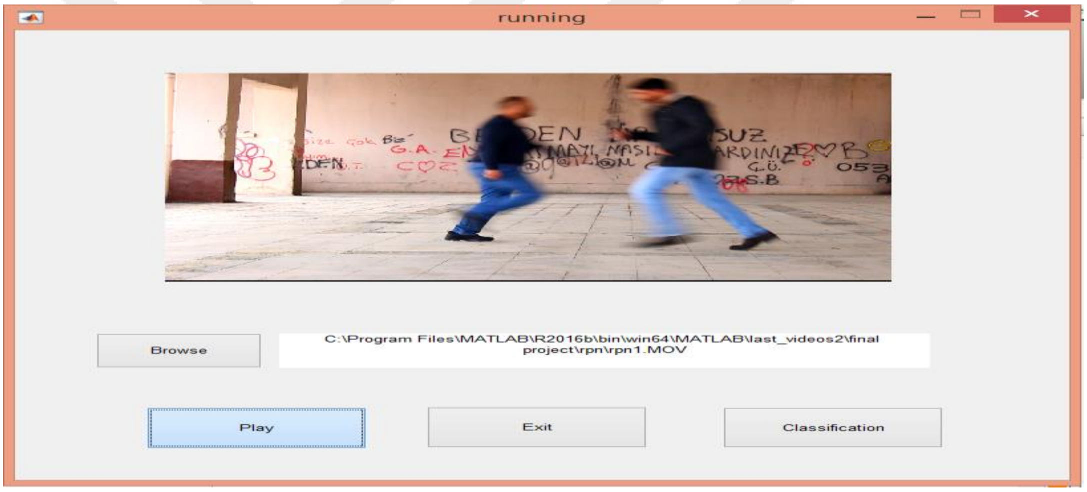


Figure 39: GUI after apply our system.

4.3 Confusion Matrix:

The matrix of confusion clarify the results that obtained and put in a matrix with the experiments that were performed to calculate the accuracy ratio and evaluation the performance, where the columns in matrix represent the expected cases for a particular category, either columns represent the actual cases corresponding for that expected cases[20] see figure (40). Where this matrix has 4 results which are:

- **True Positive "TP":** refers to positive statuses that identified by the system correctly.
- **False Positive "FP":** Refers to negative statuses that identified by the system incorrectly as positive.
- **False Negative "FN":** Refers to positive statuses that identified by the system incorrectly as negative.
- **True Negative "TN":** Refers to negative statuses that identified by the system correctly.

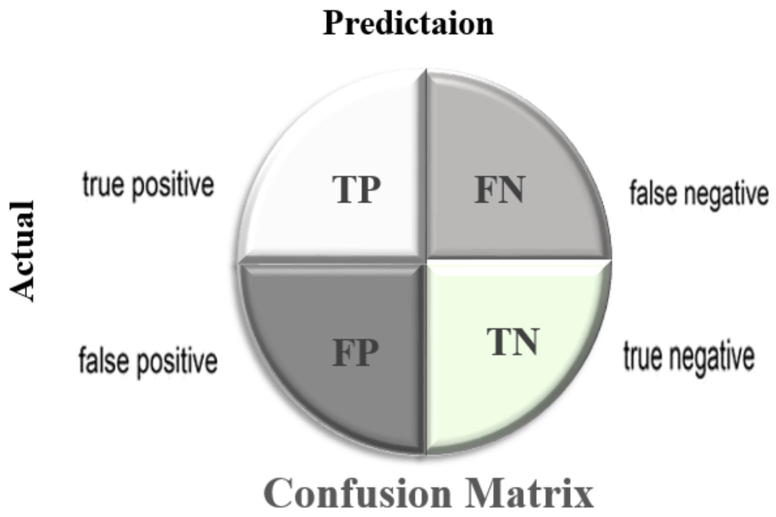


Figure 40: Confusion Matrix.

CHAPTER 5

CONCLUSION AND FUTURE WORK

This proposed system addressed the suspicious activities indoor by fixed camera all movies were in MOV format all videos consists of 23 frames per second., where we discussed most cases of running and walking (parallel with the camera" right to left/ left to right", To/away from the camera), we applied all previous cases on one and two objects that was by using background subtraction algorithm, thresholding and used blob analyse for track the object, we have faced many challenges such as reflection, shadows and lighting, these problems have been eliminated by using morphological operations like dilation and erosion. The experimental results show us that our system gives good results in detecting suspicious activities (running) and distinguish them from suspicious activities (walking).Where the results gave us a huge difference between the different cases, for parallel cases wıth the camera or in cases of (to or away from the camera), we note that the X-axis relative to the cases of the parallel with camera increases or decreases depending on the direction of the object also note that the size of the object varies depending on the location of the object of the camera. So by all these differences we can know if this activity is suspicious or non-suspicious.

FUTURE WORK

In video surveillance field there are open area for researchers, we tried to cover all cases of running and walking activities and computed speed and size of all objects in the movies to detect the suspicious/non-suspicious activities. The following points can be considered as supplementary research for this research:

1. Cover more suspicious activities such as fighting, jumping...etc.

2. Capture the movies by using moving cameras instead of fixed cameras, also from different sides, furthermore can be use more than one camera to capture the movie at same time.
3. The possibility of connecting the movies with a mobile phone application and trigger alarm or send warning messages to security agencies in case of detected suspicious activities.
4. It can be extended the tracking part by using "Kalman filter" to provide track of each object in frame.

REFERENCES:

1. **N. Ihaddadene and C. Djeraba., (2008),** "*Real-time crowd motion analysis",* IEEE, pp 1-4.

2. **K. Yokoi, T. Watanabe, and S. Ito (2009), "**Surveillance *Event Detection Task"*, pp 1-10.

3. **W. Hu, T. Tan, L. Wang, and S. Maybank (2004),** "*A Survey on Visual Surveillance of Object Motion and Behaviours"*. IEEE, Vol: 34, Issue 3, pp 334 – 352.

4. **Divya J, M.E, G. H, Pune, Prof. Dr. R.S.Bichkar (2015),**" *Automatic Video Based Surveillance System for Abnormal Behaviour Detection".* IJSR, Vol: 4 Issue: 7, pp 1743- 1747.

5. **Kaushik Deb, Sayem Imtiaz, and Priyam Biswas (2014),**" *A Motion Region Detection and Tracking Method*", vol. 4, no. 1, pp79-90.

6. **Prof. D. D. Dighe, Ms. K. V. Patil (2014),** "Background Subtraction Algorithm Based Human Behaviour Detection ",Vol. 4, Issue 7(Version 3), pp.60-64,ijera.

7. **Srinivasa Rao Chalamala, Nisha Pal, Suneeta Agarwal (2015),** "*Detection and Tracking of Moving Objects in Surveillance System".*Volume.14, pp 2321-1814, ISBN (Print): 978-1-62951-946-3.

8. **R.Naveen Kumar and Dr.S.Chandrakala (2016),** *"DETECTING AGGRESSIVE HUMAN BEHAVIOR IN PUBLIC ENVIRONMENTS "*Department of computer science and engineer, Rajalakshmi Engineering College Chennai, Tamil nadu. ISSN: 0976-1353 Volume 22 Issue 2.

9. **Pooja N S., (2015),** "*Suspicious Group Event Detection for Outdoor Environment*", (IJMTER), Volume 0X, Issue 0Y, pp 611-618.

10. **C.Srinivas Rao, P.Darwin (2012),** "Frame Difference and Kalman Filter Techniques for Detection of Moving Vehicles in Video Surveillance", Vol. 2, Issue 6, pp.1168-1170, (IJERA).

11. **D STALIN ALEX, Dr. AMITABH WAHI, (2014),** *"BSFD: BACKGROUND SUBTRACTION FRAME DIFFERENCE ALGORITHM FOR MOVING OBJECT DETECTION AND EXTRACTION"*, Vol. 60 No.3, JATIT & LLS.

12. **Christopher M. Bishop, M. Jordan, J.Kleinberg, B. Schölkopf (2006)**, *"Pattern Recognition and Machine Learning"* Springer.

13. **Yung-Tai Hsu, Hong-Yuan Mark Liao, Jun-Wei Hsieh, Chih-Chiang Chen. (2008),** "*Video-Based Human Movement Analysis and Its Application to Surveillance Systems*", IEEE, VOL. 10, Issue: 3, pp.372-384.

14. **Chitra Hegde, Shakti Singh Chundawat, Divya S N (2016),** "*Unusual Event Detection Using Mean Feature Point Matching Algorithm"*.Vol.6, pp.1595-1601,ISSN: 2088-8708, DOI: 10.11591/ijece.v6i4.10179.

15. **R.T. Collins and et al (2000),** "*A system for video surveillance and monitoring"*. Technical Report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA.

16. **Mohamed Elhoseiny, Amr Bakry, Ahmed Elgammal (2013),** "*Multiclass Object Classification in Video Surveillance Systems Experimental Study",* IEEE, USA. Vol 2, pp. 61–66.

17. **Srinivasa Rao Chalamala and Prasanna Kumar (2016),** *"A Probabilistic Approach for Human Action Recognition using Motion Trajectories",* IEEE, India, pp 185-190.
18. **Amel Ben Mahjoub and Mohamed Atri (2016), "***Human action recognition using RGB data*"Tunisia, Monastir University, IEEE, Tunisia, pp 83-87.

19. **Yang Li, Qin Lu, Wusheng Luo (2016),** "*Adaptive temporal sampling for real-time human action recognition",* College of Mechatronics Engineering and Automation National University of Defense Technology Changsha, China, IEEE, pp 1306-1310.

20. https://www.scribd.com/document/317546158/Confusion-Matrix-based-Feature-Selection.

21. https://www.mathworks.com/help/vision/examples/detecting-cars-using-gaussian-mixture-models.html.

22. **S. Y. Elhabian, K. M. E1-Sayed, S. H. Ahmed (2008),** "*Moving object detection in spatial domain using background removal techniques- State-of-Art*", Recent Patents on Computer Science, vol. 1, no.1.

23. **Naveen Javed, O. & Shah, M. (2002),** "*Tracking and Object Classification for Automate Surveillance*," Proceedings of the 7th European Conference on Computer Vision, Part-IV, pp. 343-357.

24. **Komal Rahangdale, Mahadev Kokate (2016),** "*EVENT DETECTION USING BACKGROUND SUBTRACTION FOR SURVEILLANCE SYSTEMS",* Vol. 1, Issue 4, ISSN No. 2455-2143, Pages 25-28.

CURRICULUM VITAE

Surname, Name: Salem, Fathia.

Date and Place of Birth: 12 July 1983, Libya.

Marital Status: Married**.**

Phone: 545 834 3701.

Email: **almnfy_fat@yahoo.com.**

EDUCATION

Degree	Institution	Year of Graduation
M.Sc.	Çankaya Unıv.,Computer Engneering.	2017
B.Sc.	Omar Al-Mukhtar Unıv., Computer Science	2005

WORK EXPERINCE.

Year	Place	Enrollment
2010	Omar Al-Mukhtar Unıv.,Department of Computer Science	Demonstrator
2007	El-Galaa Secondary School	Teacher

FORGEIN LANGUAGES

English.

HOBBIES

Reading, Travel.