Design of an ATM Security through Smart Vision

DISSERTATION-II

Submitted in partial fulfillment of the

Requirement for the award of the

Degree of

MASTER OF TECHNOLOGY
IN
Electronics & Communication Engineering

By

(Bharti Thakur)

Under the Guidance of

Prof. Bhupinder Verma

School of Electronics and Electrical Engineering (SEEE)

Lovely Professional University

Phagwara, PUNJAB

December 2017

CERTIFICATE

This is to certify that the dissertation titled **"Design of an ATM Security through Smart Vision"** that is being submitted by Bharti Thakur is in partial fulfillment of the requirement for award of **MASTER OF TECHONOLGY** is a record of bona-fide work done under my guidance. The contests of this dissertation, in full or in parts, have neither been taken from any other source nor have been submitted to any other Institute or University for award of any degree or diploma and the same is certified .

Prof. Bhupinder Verma

Professor and Head,

School of Electronics and Electrical Engineering

(Lovely Professional university)

Examiner (I) Examiner (II)

ABSTRACT

A vibrant branch of research in computer vision that has attracted a lot of attention for decades is the human activity understanding from video. A means for accurately locating humans in image or a video is a prerequisite to the process of understanding human activities or action. This work's focus is on investigating the use of people detectors for video surveillance in Financial Banks premises so that it can eventually be used for abnormal human activity detection. Posture recognition is one of the most interesting fields in computer vision because of its numerous applications in various fields.

ACKNOWLEDGEMENT

Foremost, I would like to express my sincere gratitude towards my guide, **Prof. Bhupinder Verma**, who gave his full support in the working of this dissertation work with his stimulating suggestions and encouragement all the time. He has always been a source of ideas.

I am thankful to management of Lovely Professional University for giving me an opportunity to carry out studies at the university and providing the proper infrastructure like internet facility which is pool of vast knowledge to work in.

I owe my heartiest thanks to my parents & all those guidepost who really acted as lightening pillars to enlighten my way throughout this project that has led to successful and satisfactory completion of my Dissertation-II.

At last but not the least I am thankful to all my friends, , who have been instrumental in creating proper, healthy and conductive environment and including new and fresh innovative ideas for me during the last date of submission of project. Without them, it would have been extremely difficult for me to prepare the project report in a time bound framework.

I would like to thank God for the strength that keeps me standing and for the hope that keeps me believing that this report would be possible.

Bharti Thakur
Registration no. 11611923

DECLARATION

I, **Bharti Thakur** student of **M.Tech (ECE)** under School of Electronics and Electrical Engineering of Lovely Professional University, Punjab, hereby declare that all the information furnished in this dissertation report is based on my own intensive research and is genuine.

Date:

Bharti Thakur

Registration No. 11611923

Certified that the above declaration made by the students is correct and best to our knowledge and belief.

Prof. BhupinderVerma

Supervisor

School of Electronics and Electrical Engineering

Lovely Professional University

LIST OF FIGURES

List of Table

List of abbreviation

HOG	Histogram of Oriented Gradient
HCI	Human Computer Interface
PCA	Principle Component Analysis
SVM	Support Vector Machines
MEI	Motion Energy Image
MHI	Motion History Image
SURF	Speed up Robust Feature

TABLE OF CONTENTS

CHAPTER 5

Chapter 1

Introduction

Modern time we can see that cameras are used everywhere, such as banks, shopping malls, campus area, residential area, institutes etc. at such places it is more important to maintain security. The main objective for movement recognizes the events and aims of one or more people from the progression of observations on the people's action. Supervised learning and understanding of unusual human being deeds is more complex, difficult and wide task. Human act Recognition and extracting feature from the make full use of and derive benefit from (a resource) make full use of and derive benefit from (a resource) make full use of and derive benefit from (a resource) database has usual a lot of consideration in the computer-vision, machine learning communities [1].

As information technology continues to evolve quickly, Human-Computer Interface (HCI) is playing a role of growing importance. Even after huge development of input devices, sill the interaction with computers is cumbersome and an uncomfortable experience. The adaption of computers to natural means of communication like speech and body language has changed the whole experience of interacting with machines. Thriving realization of these modalities into an edge has the possible of easing the HCI bottleneck that has become noticeable with the advances in computing and communication. The HCI can be briefly explained by the point of communication between the human user and the computer.

We usually interrelate with computers using devices. Computer devices become vaster and provide the user friendly interface with computers. Embedded system is used almost in every walk of our day to day life; we need more intuitive means of interacting with computers. Currently HCI devices are not intuitive or common man. It needs some training before using those, means first we train the system and also they limit application of computers in daily life. As they need the presence of human beings in the surrounding area of computers. Since posture form very important part of our communication. It is designed to recognize human gestures and can provide more efficient in taking computer tools approximate to human being. These results with which human communicate with each other are integrated into HCI with an aim to determine the efficient utilization of the available information flow of the communication, and demonstrate technologies.

In current days, there has been a remarkable significance to introduce the intuitive interfaces that can identify by the user's body movements and transform them into machine level language. Surveillance is the used to observe the behavior, changing information, or activities, usually of people for the cause of influencing, organization, and directing or to defend them. Surveillance is used for various works such as government for intelligence gathering, the prohibition of crime, the protection of a process, person, group or object, or for the inspection of the crime scenes. A useful video surveillance system depends on detection of suspicious activities [2]. In recent times, detecting the abnormalities in human behavior using such gained more substance as it can provide clues while preventing breaches in security.

Human behavioral pattern and face recognition plays vital role in person identification. Visual information is a key source for such identifications. Surveillance videos provide such visual information which can be viewed as live videos, or it can be played back for future references. Now days 'automation' has its impact even in the field of video analytics. Video analyzing can be used for wide variety of applications like human activity prediction, person identification, motion detection, abnormality prediction, people counting at crowded places, vehicle counting etc[2]. The actions of human may be overlapped by background changes or camera motion. Especially, camera motion may interfere with detecting the true human activity. Secondly, it is still challenging to recognize human activities in realistic unconstrained videos because of a large amount of intra-class action variations. we consider the challenging problems mentioned on top of, and focus on videos frameworks and developing a model for detecting abnormal activity using HOG features and SVM classifier.

There are two stages in this architecture, training and testing. In training stage we consider normal and abnormal images as input sample mages, sample images which we are using as input sample are taken from the videos which we are testing and later those images are pre processed from color images to gray scale images. To foreshorten human activity accurately and efficiently, the HOG features are employed and PCA for feature reduction, result from feature reduction are trained to SVM classifier and will dumped to knowledge base as a result database.

In this phase we train the SVM classifier to classify the human activity on the basis of result. In testing phase we consider human activity video as input and generate the frames of that video, pre process the video and to foreshorten human activity accurately and efficiently, the HOG features are employed and PCA for feature reduction, SVM classifiers compares the images in

knowledge base with the human activity in the frames of videos and returns the result according to the human activity.

1.1 Objective and Goals

The goal set for this work is full body pose and human irregularity recognition for interaction where the user absorbed into an implicit world. The work presented in this aims to enhance the human computer interaction based on computer vision. It aims to improve the interaction with pioneering real –time and indicator less abnormalities and body movement-based systems. The systems described in this aspire to attain this real-time. They are built with normal computers prepared with typical tint cameras, which are often even found, built into the displays. Indicator less system allow usual interface without wearing special markers or special tracking suits, which are generally required in modern day tracking systems used in the entertainment industry. Concurrent systems have a high revive rate and minimal latency, providing the user with smooth and instantaneous interaction with the system.

The present advances in computing machinery push the interest in other ways instead of keyboard, mouse, or keypads etc. human makes various gestures to communicate with other. To detect the abnormality by the movement of the body parts. Wherever tracking each and every part of the human body is very difficult because of time and space constraints, person occlusion and other impenetrable factors. The current abnormality analysis typically try to classify a set of posture made using specific body parts and try to recognize them. It can be performed by using static images and it can also be performed by using real-time. There has been vital research dedicated to head, eyes, orifice, face expressions, but comparatively small in recognizing huge pose. By using this, we track the full-body of the human being and recognitizing the abnormality.

Technological advancements in computational power have grown manifold and changes in society have roused an important augment in the figure of surveillance cameras; whether it is for monitoring traffic or CCTV security. This problem has inspired many computer vision researchers to develop intelligent means for machines to recognize objects and actions in real-time, alleviating the cognitive load on those who monitor the surveillance, highlighting potentially interesting goings-on as and when they are occurring. This project sets out to explore a number of proven techniques used for representing human actions on images datasets by extending and applying

them to realistic scenes in real-time. The outcome is a collection of results representative of the justified approaches to solving the action recognition problem.

1.2 System Design and Work

The thesis deals with the description of algorithm and methods to build vision based marker less, abnormality recognition and human computer interfacing system. The design of HCI tool based on illustration inputs consists of a numeral of steps. Also that the project is conducted over a relatively modest era of time, progress is essential to be quick, effective and flexible to change in order to successfully achieve all objectives. Essentially, the design is interleaved with the actual codification of the system allowing more time to be spent on testing if it works. This mentality is suitable for this particular projects as there is a clear approach to producing a solution to the problem, while the considered techniques are all proven leaving much of the effort committed to justification, implementation and testing, followed by experimentation and evaluation.

The system considered here is vision-based. The vision based means that the system uses one or more cameras to observe the user, or objects that are attached to the user. the advantage of these system is that the user does not need to hold or touch the device with which he is interacting, which is especially useful if the user is interacting with large displays or is immersed in virtual environments.

Non-vision based devices include touch screens, exoskeleton, suits with sensors, gloves with sensors, and accelerometer-based remotes. The system is marker less and so the user does not need to wear special suits, helmets, gloves, making the system more natural and intuitive. Furthermore, the hardware requirements can be made smaller, as the system can use standard cameras which are already present in most recent systems.

Research in the field of marker- less vision based systems is ongoing and existing systems are very experimental. Both offline and real-time systems is evaluated for the work. Offline means that the images are recorded and then the tracking algorithm is run offline. These systems allow for accurate tracking of the person remains, which can be helpful for analysis of the persons in the video, motion capturing, or archiving video content. When the user expects real-time interaction with the system then offline mode is not used. Also real-time interaction provides with more applications like gaming, security and comfort.

1.3 Scope of Study

Abnormal activity detection is very interesting field and it has got interesting future works, we have seen different kind algorithms like HOG, SURF, Harris corner detection, fast method, Graph cut method and many more which are implemented in different ways to overcome the possible problems occurring in detecting abnormality in ATM surveillance. In this for feature detection in images computation time is quite high .Computational complexity is more in most of the algorithms which can be improved further. So one way is that the computation time for the feature detection from images can be reduced and also it a possibility that work can be focused on more geometric distortion.

Work can also be done in order to ease the cumulate errors when edging the large quantity of images and calculate images focus more accurately. The arrangement of the detectors varies according to the database images. Hierarchical classifiers foliage could provide the vital information about the preparation of databases and for betterment. Human Body parts provide information which deserves better interest and study to cope with occlusion and recognize abnormality in composite surroundings.

1.4 Literature Review

Innumerable attempts are done to detect and recognize the human actions occurring within the video sequences have been made. Davis and Bobik [3] present a new approach to their presentation and recognition of human movement. They consider video sequences, computing sequential templates constituting two mechanisms; a MEI and a MHI. The MEI is a binary illustration of the motion occurring between sequences of frames while the MHI is a grayscale intensity depiction where the most recent changes are lightest. Provided that the action can be located and extracted from a busier scene, this temporal action representation is useful as it captures the method in which the figure and motion develop over time.. In spite of the approximately total lack of familiar features in the motionless imagery, the faction is easily predictable when the series is given to the stroke screen. This capability of the human vision system assert for recognition for recognition of movement directly from the motion itself, as incompatible to primary reconstructing a 3-D model of a person and then recognitizing the motion

of the model. Firstly the proposed a representation and recognition theory that decomposes action based recognition into first recitation where there is motion , the spatial pattern and then explain how the action is stirring.

There are strategies to be followed in order to locate the things of interest example. A human, tracking it so that a description how the object changes over time can be formed, and then finally classifying the stroke. Stauffer and Grimson [4] used this approach to recognize general activities by modeling the changes in velocity, position, size among other quantitative features. One the difficulty in tracing is minimizing the errors which occur when foreground segmentation is poor. Tracing through motion or foreground segmentation can be sensitive and degrade ungracefully when errors occur in adjacent frames within a capture sequence [5].

Human action recognition solutions must be able to detect the region where an action is done within the video sequence before it can be classified. This is opposed to an action specific search which would involve considering a given action and seeing if it is arise within a video. There are numerous ways in which interesting regions potentially featuring action could be determined. One such technique is built around Harris Corner detector [6], a popular operator used within the object recognition to detect key points which contain high information.

Actions can be represented such that they are suitable and easily understandable for Machine Learning Algorithms using two techniques. The first technique is Histogram of Oriented Gradient (HOG) which derives from the object recognition problem due to its quality in describing shape. It is used for human action recognition is inspired by its successful application to pedestrian detection by Dalal and Triggs[7]. If the shape of human changes over the temporal dimension of the space-time cuboids is describes by HOG. There is other technique like HOG i.e. Histogram Optical Flow (HOF) [8]. This is a more natural, flexible and effective method. The idea of optical flow within action recognition assumes a fixed position camera such that the impression of motion can be concluded. While HOG simply requires the spatial gradient derivatives, optical flow estimates can be computed using a number of techniques, namely those by Lucas and Kanade [9] and Horn Schunck [10].

Human behavior and action for identification or uncovering of unusual event has fascinated by vital research interest in current years. It is not only updates of earlier associated surveys, but also a focus on abnormal human actions detection particularly video supervision applications. There are so many public places where need to keep track of activities and need to put security system

for protecting such public places. In this paper recognition of abnormal human behavior involves various modeling and categorization of abnormal human behavior with assured rules. There is already research is done in this field of vision based abnormal behavior detection but every approach has its own advantage and disadvantage. There are so numerous methods to detect irregularity such as Hidden Markov Model, Bayesian Model, Support Vector Machines, Neural Network, Fuzzy, etc. HMM rule is semi-supervised learning rule which require normal event recognition form a huge dataset and abnormal events are learned by Bayesian adaption in an unsupervised manner [11].

Kalman filter is usually used to track objects. HOG features for being and used for oblique mapping based upon that maps some detected visual features to the number of people [12]. If someone enters in the ATM and by doing normal behavior leaves the ATM. Firstly, capture is separated into frames. Magnitude is calculated by root of sum squares of individual pixel concentration in grey level over frames.HOG is applied to extract information of the matrix [13].

Histogram Oriented Gradient use ably describes the shape and outer shell of local object with in an image by share of intensity gradients and edge direction. This the way how to dataset is generated of the training video clips. Author used a method ‘Random Forest’ for classification. This method is trained by some specific rule. It is strong enough to remove noise in training dataset. Conclusion, there are so many methods for abnormality detection such Histogram of Gradients, Kalman filter, Histogram of Optical flow, Spatial temporal etc. Spatial temporal features can recognize abnormal activities better than other methods [14].

Posture identification is one of the most elegant fields in the computer vision because of its numerous applications in various fields. We can solve the various problems by using 3-dimensional camera instead of using static cameras. In this author discover a technique of using skeleton information, provided by Kinect 3D camera . It is used for Posture recognition or efficient real-time ATM intelligent monitoring. To get the desired results posture recognition, Kinect camera is used to track the joints of bone in the and their position in sequence, the system detects the unbalanced behaviors. It is divided into three steps: data acquisition where we acquire the data from the image, data processing where we apply various image processing operations and feature extraction and posture identification.

To approximate the likelihood of posture being anomalous depends on the skeleton information. Regression study is a good approach to solve this problem. Three modules are used, first data acquisition, using kinect, we confine different types of information that is color, depth, and

skeleton in order. Second step is data processing, which deals to finding the 3-D locations of body-joints. It is using OpenNI framework. Human posture identification deals with the learning of posture using as a training dataset and classifying the test posture using one of the predefined classes: if abnormal activity is find then alarm is raised ,there is no alarming condition for normal activity. In data processing the user position data can be accessed which provides information about various joints in the skeleton. In this it tracked the joints, if joint is tracked it set the status 'tracked', if the joints are not tracked than it will check the other joints, and recognitizing the abnormality. Posture recognition can be achieved through logistic regression, a machine learning technique that for a given input predicts a class and provides a probability associated with the prediction. These possibilities are very important, as they provide more precise output in predictions of anomaly detection [15].

Currently study is going on in the field of crime recognition and avoidance in the ATM. But till now there is no advanced technology came in the field of ATM to grant security. In this paper the first part comprise on video camera which capture the images. Second part is numerous object recognition modules which detect the survival of more than one person in ATM room. If it is normal interaction then it passed to usual transaction module and transaction take place. If there is multiple objects and consumer not allowed multiple objects, then it will produce alarm and make call to nearest police station or security office. Information flow in proposed intelligent agent as follows: when customers enter into the ATM room, his image will capture by camera. When a customer started transaction, then the control and information flow in proposed system. Anomalous activity recognition in capture has many vital computer vision applications, such as video supervision, HCI, video browsing, and analysis of human actions etc. the anomaly recognition can be used to alert the related ability of potential criminal,theif or hazardous behaviors.

For identifying the movement three major steps of processing are considered, that is, object segmentation, feature extraction and illustration, action recognition and classification algorithms. The first step is to segment a person from the capture series. The second step is extracted the features from the database, such as shape, edges, mass, actions etc. The last step is to recognize the abnormal behavior. There are a variety of methods used for recognizing the person action in computer vision [16].

The human action detection is a vast and currently used area of research and its associated study society not only because of its potential benefits but also due to the numerous capable

applications that it is a vital part of them. To detect people in descriptions and videos is one of the important challenges in computer vision. To find the abnormality is a vat area of research and currently it is using, identify all human in image. In this two techniques are used HOG and Haar integral, HOG is used to extracting the features and Haar integral is used to locate the human. In this framework it includes: preprocessing, detector framework, and the post-processing module. The software used is Matlab. It provides an in build library package support toolbox in the Computer vision System using matlab.

HOG is used for detecting object in images, by using HOG descriptor features of images are extracted and trained the system.HOG descriptor is divided into blocks 4x4,8x8,16x16 pixels. First is aspect to extract the features from the images but the descriptor is used for the extraction of features from the image must be a good descriptor. Categorization plays vital role in machine learning. Second, Haar- likes features. It is used to detect the upper-body of the human being. Abnormality recognition required various methods so we get various boundary boxes, we merge the boundary boxes and at last post processing is done to remove the extra bounding box from the images. For post processing there are so many techniques or methods, we apply various methods to remove the boundary boxes from the image and last we get the desired output.

In the we compare the various boundary boxes and remove those boundary boxes those are overlap. It is shown in the paper that it is not an appropriate method to detect the abnormality because if the person not captured in the camera it provide problem for the recognition of the abnormality. When person not facing the camera, it become difficult to detect the natives. The main perpose of this is to detect the abnormality and capture supervision that integrating the two methods or techniques; HOG descriptors for the feature extraction and Haar like vital image for upper part of the body detection [17].

In this paper real time video surveillance is used for detecting the abnormal activity. In this two approaches are used first is PCA and SVM is used for classification of person behaviors. The second is to find the velocity of the pixels. Both the algorithms are effectively implemented in swarming environments for detecting the human's unusual behavior. Various posture are used for irregular activity, such as 1) running people in crowded area. 2) People are walking or standing. 3) Moving hands up and down. Deeds analysis is divided in unlike steps, first is preprocessing in this foreground detection, segmentation, and tracking the system collects continue blobs for a single person. Second PCA is used to extract the features of the images and trained the system and provide testing of the system. By using PCA approach we can collect good amount of features,

which is used for the testing part. Third is SVM classifiers, it is used to separate the behaviors of human detection such as normal or abnormal posture. The SVM is a latest method in the machine learning.

It described a concurrent capture supervision system for detecting the human anomalous behavior. To identify anomalous behaviors two approaches are used. The initial one employs PCA to take out the features from the training image and second is SVM, which employs two module 1) normal behavior 2) unusual behavior. Anomalous behaviors: bending person, running people. Normal behaviors: standing people [18].

In this, the first step is for usual actions identification uses surveillance of human activities. Abnormal activity is recogonitize by using surveillance cameras. A list of significant actions identification paper is accessible, including behaviors, datasets, realization techniques, and results [19].

In this, described a fitting algorithm for self-government human-movement recognition for video supervision. In proposing system extract human-movement recognition and tracking natives in real-time using Gaussian Mixture Model (GMM) [20].

1.5 System Overview

The work is based on the abnormality recognition using input images for the analysis of abnormality. It includes the main stages that are background subtraction, feature extraction and abnormality recognition. Each represents an active research area in itself. Background subtraction is done when we applied it to real-time applications than Background subtraction is used. After the stage of image preprocessing which may include image de-noising, post processing like morphology operations etc. Morphological operations like Dilation, Erosion, Median filter, canny edge detection. Median filter is used to remove the noise from the image. To detect the edges we use canny method.

Background subtraction is used for various applications such as abnormal activity detection, gesture recognition, and face recognition methods. Background subtraction is used when we are working in real-time applications. For background subtraction we are using various methods. It includes various steps, first we subtract the original image from the operated image than we convert the image into RGB to grayscale and then we convert the image into binary

image. After that we apply median filter to remove the noise and the we apply thresholding and subtract the background.

Feature extraction technique is used to extract the features of the images. We applied various techniques for feature extractions such as HOG feature, SURF etc. first we trained the system and then provide the various other techniques to find the abnormality in an ATM surveillance. When it comes to study of the postures then it is just unfeasible to make an entire set of postures because human postures are infinite and they construct the use of each and every part of the body. When recognizing using images, I have considered some postures that are standing straight, moving left, moving right, hands up, hands down.

1.6 Scope of the Study

Abnormal activity detection is very interesting field and it has got interesting future works, we have seen different kind algorithms like HOG, SURF, Harris corner detection, fast method, Graph cut method and many more which are implemented in different ways to overcome the possible problems occurring in detecting abnormality in ATM surveillance. In this for feature detection in images computation time is quite high .Computational complexity is more in most of the algorithms which can be improved further. It is a wide area of research to detect the people and find the abnormality in the real life. It is more advantageous or it can be used in various applications in real-time such as security: because cameras are used widely, bank surveillance, ATM security, jewelers shop, treasure security etc.

There are various methods used to detect the abnormality. Random Forest method, SVM, Decision tree, Bayesian method, K-mean clustering etc. SVM is preferred because it is more advantageous, it is used in real-time applications. It is used for gesture recognition, face recognition, finger print recognition etc.

1.7 Organization of the Dissertation

This present the form of individual building blocks that can be used to build interaction systems. The building blocks can be separated into two parts: (1) preparing the input in support of recognition and identification (it includes segmentation) and (2) features extraction. These building blocks serve as subsystems for the posture recognition system.

Chapter 2 Proposed future work Plans and Timelines.

Chapter 3 describes the different segmentation steps that are performed on input images for tracking the object i.e. the human body. Background subtraction, foreground-background segmentation separates the user from the background. describes how gestures are recognized using support vector machine. deals with action representation of the body using different feature extraction techniques like HOG. Actions can be described using one of the features; shape, motion or mass.

Chapter 4 Research methodology or working steps for the project.

Chapter 5. Provides the results, conclusion and Future Prospects of the works presented in Dissertation-II..

Chapter 2

Proposed future work plans and Timelines

Table2.1 proposed work plans and timelines

work	Jan-March 2017	April-June 2017	July-Sep. 2017	Oct.-Dec. 2017
Literature review completed				
Base-paper selection				
Preparation for database				
Implementation Of the base paper				
Result				
Report				

In this we proposed the work plans, literature review is completed between the month of the august and September. After the completion of the literature review, the second step is for the selection of the base paper. For the implementation of the base paper we required database. Then

the third step is for the preparation of the database and then the implementation of the base paper. After the implementation of the base paper, last step is for the report writing.

Table **2.2** future work plans and Timelines

Work	**Jan 2018**	**Feb 2018**	**March 2018**	**April 2018**
Database of Image				
Background Subtraction				
HOG Features + SVM Classifier				
Abnormal Activity Recognition				
Report				

Chapter 3

Image processing operations

3.1 RGB to Grayscale

RGB (Red, Green, and Blue) refers to a method for on behalf of the colors to be used on a computer exhibit. Red, green, and blue can be collective in a variety of extent to get hold of any color in the detectable spectrum. Levels of R, G, and B can each range from 0 to 100% of complete concentration.

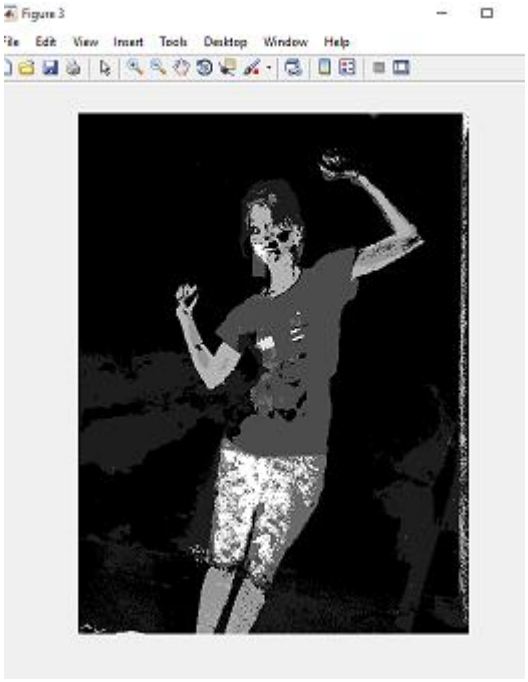


Figure3.1: RGB to Grayscale

Grayscale metaphors are different from one-bit bi-tonal binary images, images consist of merely two colors black and white also known as binary images. Grayscale images have various colors of gray in among. Grayscale images can be the effect of measuring the intensity of light at each pixel according to a fussy prejudiced combination of frequencies or wavelengths, and in such

cases they are monochromatic, accurate when only a distinct frequency (in practice, a contracted band of frequencies) is captured. The frequencies can in principle be from somewhere in the electromagnet band, for e.g. Infrared, visible light, ultraviolet, etc..

3.2 Grayscale to Binary image conversion

A black and white image is a digital image that has merely two probable principles for every pixel. In general, black and white are two colors used in the binary image. The color used for the object(s) is shown in figure is the forefront color whereas the respite of the image is the backdrop shade.

Figure3.2: Grayscale to Binary

3.3 Edge Detection

Canny edge recognition is a method to extract valuable structural information from dissimilar vision objects and significantly reduce the quantity of data to be processed. It has been extensively applied in a variety of computer vision systems. Canny has established that the chuck for the purpose of edge detection on diverse vision systems is comparatively similar. Thus, edge detection is a key to deal with these chuck can be implemented in a spacious collection of situations.

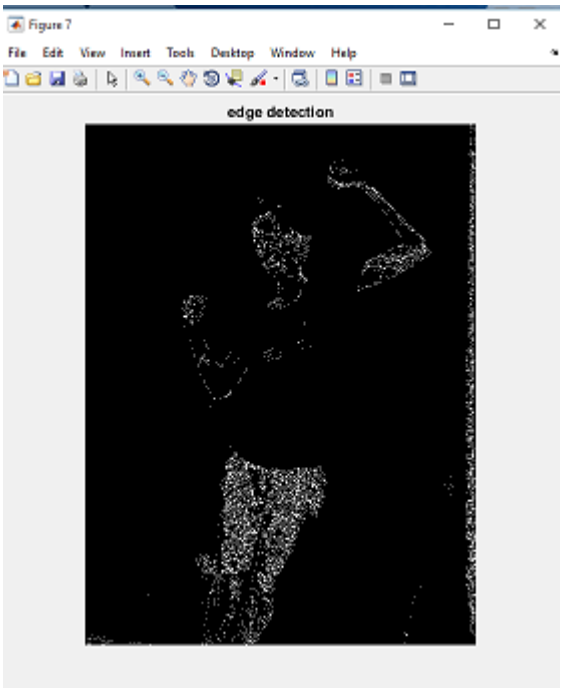


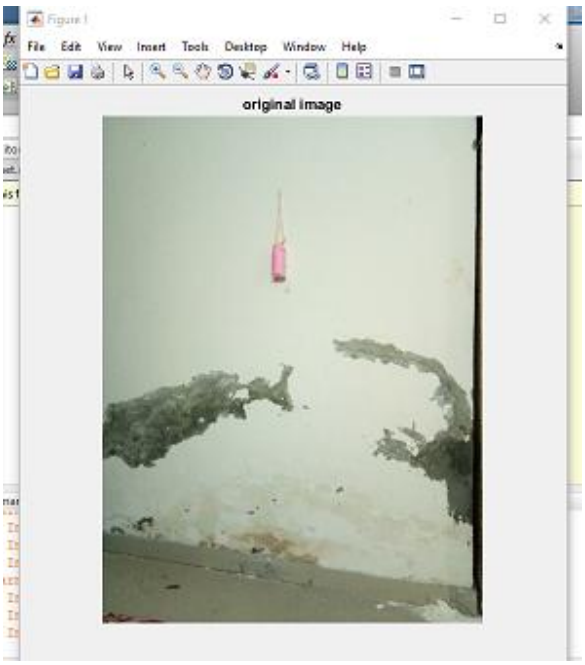
Figure3.3: Edge Detection

The main stages for edge recognition consist of:

- Detection of edge with small error rate, which means that the recognition should precisely catch as numerous edges shown in the image as probable
- The edge points can be detected from the operator should accurately limit on the center of the edge.
- A particular border in the image should only be discernible once, and where probable, image noise should not create forged edges.

3.4 Background Subtraction

With computers being embedded in each walk of our life, there is a growing demand for spontaneous devices for HCI.. As human beings use expressions as vital means of communication, devices based on anomaly detection systems will be effective for human interaction with computers. Yet, it is extremely vital to keep such as non-intrusive as possible, to reduce the limitations of interactions. To design in such a way, spontaneous, camera based real-time abnormality identification method has been a vigorous as the area of research in the meadow of computer vision.



(a) Original image

b) operated image

(c) Background subtraction

Figure3.4: a) original image b) operated image c) background subtraction

3.5 Histogram of Oriented Gradient (HOG)

HOG is a characteristic descriptor. It is used to make the classification easier under divergent conditions and characteristic is to establish the thing in such a mode that the same object produces as close as possible to the same feature descriptor. The developers of this approach trained a SVM to identify the HOG descriptors of natives Abnormal activity detection is very interesting field and it has got interesting future works, we have seen different kind algorithms like HOG, SURF, Harris corner detection, fast method, Graph cut method and many more which are implemented in different ways to overcome the possible problems occurring in detecting abnormality in ATM surveillance. In this for feature detection in images computation time is quite high .Computational complexity is more in most of the algorithms which can be improved further.

Figure83.5.1: Hands up

It is a wide area of research to detect the people and find the abnormality in the real life. It is more advantageous or it can be used in various applications in real-time such as security: because cameras are used widely, bank surveillance, ATM security, jewelers shop, treasure security etc.
There are various methods used to detect the abnormality. Random Forest method, SVM, Decision tree, Bayesian method, K-mean clustering etc

It requires a sliding recognition window which is moved in the order of the image. A HOG descriptor used a sliding window for the extraction of the features from the images. This descriptor is then deployed to the accomplished SVM, it is used as a classifier to distinguish between the classes . The image is sub sampled to multiple sizes in order to recognize persons at different scales.

Figure3.5.2: HOG 150 x150(Hands up)

Figure3.5.3: standing toward left side and keeping hand up

Figure3.5.4: HOG 150x150(standing toward left side and keeping hand up

3.6 Feature Extraction

Feature extraction includes the sinking quantity of aid vital to define a huge set of figures. When performing determination of intricate figures one of the most vital troubles caused as of the numeral of variables implicated. Analysis with a large number of variables usually requires a huge

sum of recollection and calculation supremacy; also it may cause a categorization process to over fit to guidance samples and usually poor to new samples. In the wide term for methods of fabricating the combinations of the variables to get around these problems while still describing the data with adequate exactness.

Figure3.6.1: Hands up

Variables - testfeatures

testfeatures

1x46656 single

	1	2	3	4	5	6	7	8	9	10
1	0.2549	0.2549	0.0741	0.1081	0.2480	0.2549	0.0848	0.0999	0.1657	0.1723
2										
3										
4										

Figure3.6..2: Extracted features for hands up

Figure3.6.3: Standing Straight

testfeatures

1x46656 single

	1	2	3	4	5	6	7	8	9
1	3.4578e-04	0	0.0131	0.1213	0.4069	0.4069	0.1081	0.0305	0.0327
2									
3									

Figure3.6.4: Extracted features for Standing Straight

Figure3.6.5: Both Hands are parallel

testfeatures

1x46656 single

	1	2	3	4	5	6	7
1	0.0113	0.0018	0.1133	0.1589	0.3820	0.1189	0.1782
2							
3							
4							

Figure13.6.6: Extracted Feature for Hands are parallel

Figure3.6.7: T-shape posture

testfeatures

1x46656 single

1	2	3	4	5	6
0.3282	0.3282	0.1887	0.1231	0.0087	0

Figure3.6.8: Extracted feature for T-shape posture

3.7 Posture Recognition

Humans express numerous of motion gestures in order to express a message to a receiver. If a computer can identify and differentiate these human motion patterns, the preferred

message can be reconstructed, and the computer can act in response accurately. In order to communicate with ocular messages to a receiver, a human expresses motion patterns. Slackly called gestures, methods are changeable but well-defined and have an associated meaning.

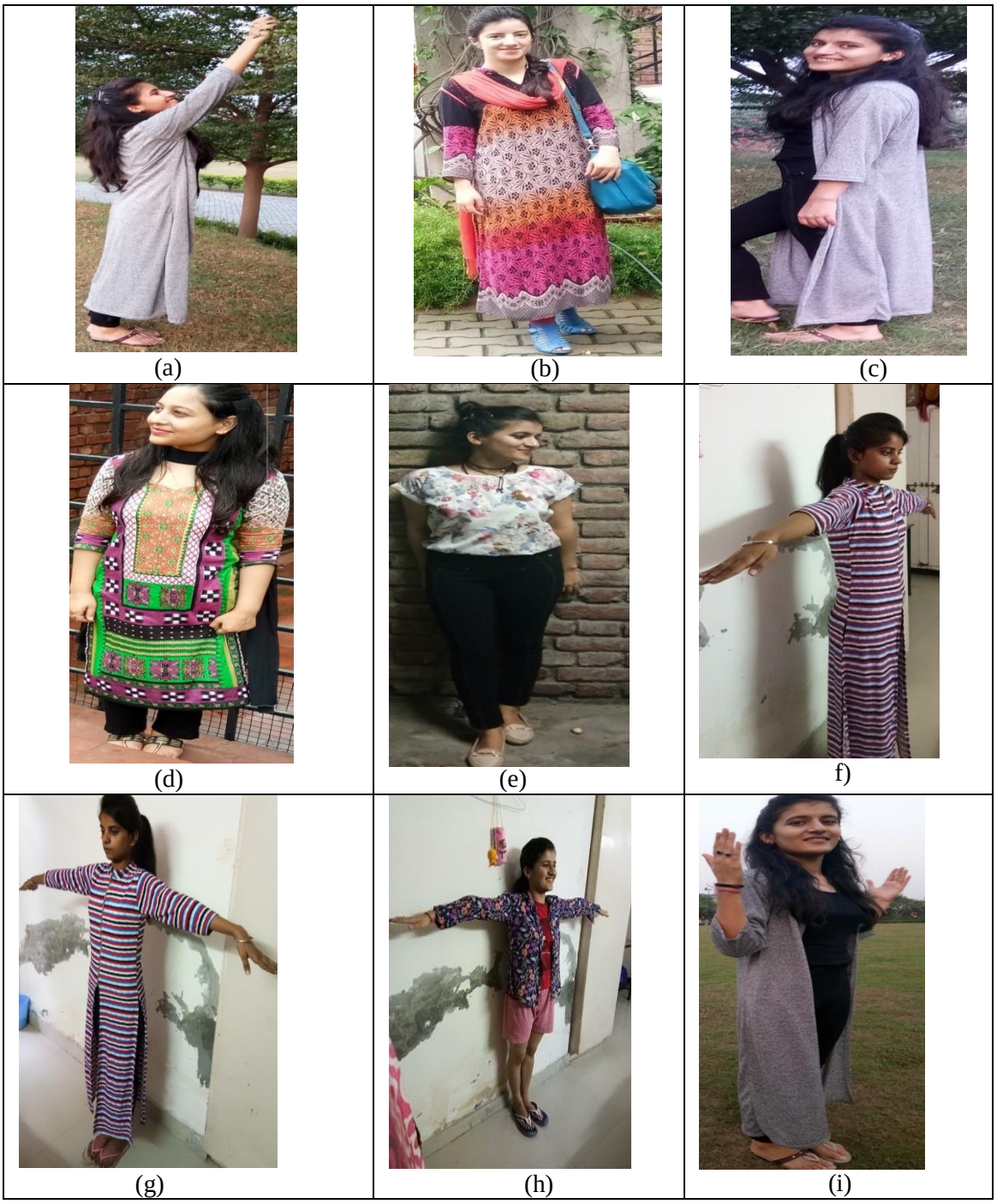


Figure3.7:Posture Recognition

Posture is explained as the movement of part of the body. If a person is moving with or without aim to mean something. Gesture forms the important feature of interaction between man-man and man-machine. HCI involves conniving a system with the aim of, to realize natural modes of message used by human beings like speaking, walking, writing, gestures etc. in their day-to-day life.

There is a numerous of Postures and to use them to convey information or for device control. Upper and lower part of the body, hands etc. used for the making of the postures. It is not possible to catalogue out all forms of postures and the system trying to identify them. SVM, in order to achieve the favorable output we divide our input into a hyper-plane. We make decision boundary or input is linear separable by a decision boundary. If we having two classes a decision boundary is drawn to distinguish between the two classes. SVM is very advantageous. It is used in real-time applications. We divide the inputs into clusters to get the output. First we trained the system and for testing we use testing images.SVM is mainly used in gesture recognition, face recognition, fingerprint recognition, gender recognition etc. a hyper plane is drawn to separate different classes.SVM is more advantageous for Human Posture Recognition, Gesture Recognition, and Face Recognition.

3.8 SVM

SVM, in order to achieve the favorable output we divide our input into a hyper-plane. We make decision boundary or input is linear separable by a decision boundary. If we having two classes a decision boundary is drawn to distinguish between the two classes. SVM is very advantageous. It is used in real-time applications. We divide the inputs into clusters to get the output. First we trained the system and for testing we use testing images.

SVM is mainly used in gesture recognition, face recognition, fingerprint recognition, gender recognition etc. a hyper plane is drawn to separate different classes.

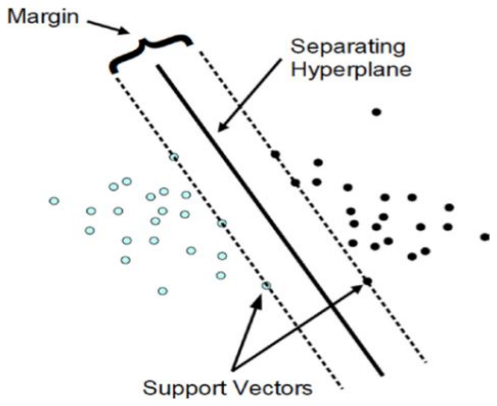


Figure3.8.1: SVM

In the below example it is shown that hyper plane separate the two classes red and blue color. A test image is used to verdict this classifier. It is supervised learning.

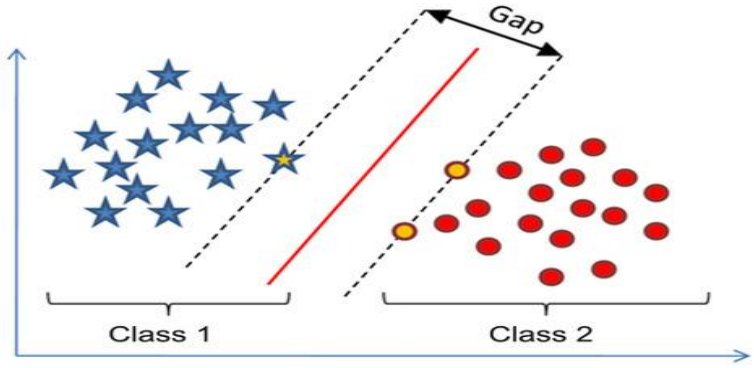


Figure3.8.2: SVM classes

A methodical example is given away in the illustration above. In this illustration, the objects fit in besides to set BLUE or RED. These colors separated by hyper plane, objects are Emerald and to the left of which all objects are Crimson. Any new object (white circle) diminishing to the right is labeled, i.e., classified, as BLUE (or classified as Crimson should it fall to the left of the sorting out line).

a) The Pros of SVM are:

- Efficacious in high dimensional spaces.

- Still efficacious in cases where number of dimensions is greater than the number of samples.
- Subdivision of preparation points in the decision function (called support vectors), so it is also reminiscence proficient.
- Adaptable: dissimilar Kernel function can be enumerated for the verdict function. General kernels are provided, but it is also possible to specify custom kernels.

b) Cons of SVM include:-

- If the numeral of features is substantially greater than the numeral of samples, evade over-fitting in choosing Kernel function and regularization expression is decisive.
- Support vector machines don't unswervingly provide likelihood estimates, these are interdentally using a high-priced five-fold cross-validation.

Chapter 4

Research Methodology

4.1 Working Steps for the Project

The simplistic illustration of structure is shown in Figure. These modalities consist of three steps:-

1) Preparation of the database images.
2) HOG feature extractions+ SVM classifier.
3) Abnormality recognition. If the abnormal activity is detected then an alarm is raised to security or either nearby police station. The software is used for the development of project is Matlab.

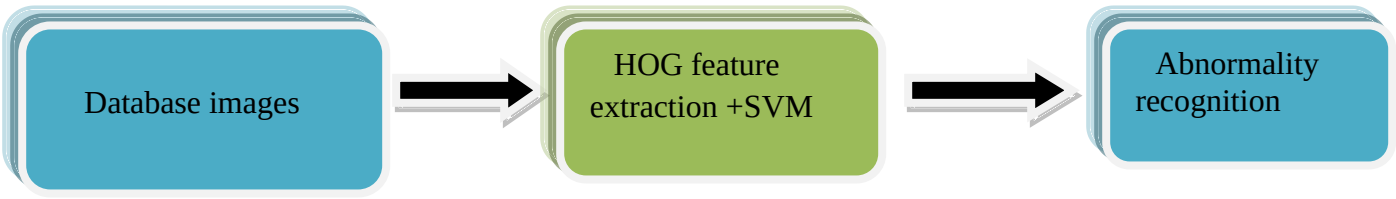


Figure4.1.1: Methodology

Gather the variety of types of images and make the database. Those images are taken as the input images. Second step, is to take out the features from the images. Characteristic extraction is generally broadly reduction.. It makes the use of algorithm to identify ad segregate a variety of preferred portions or silhouette (features) of image. HOG is used for detecting object in images, by using HOG descriptor features of images are extracted and trained the system.HOG descriptor is divided into blocks 4x4,8x8,16x16 pixels. First is aspect to extract the features from the images but the descriptor is used for the extraction of features from the image must be a good descriptor. Categorization plays vital role in machine learning.. It is used to detect the upper-body of the human being.

There are two proven features that have been shown to work in action identification;the histogram of oriented gradient (HOG). These feature representation are chosen to independently to

describe the target actions in this project as they model the changes in shape and motion over the time. Modeling the changes in a motion is an intuitive approach to action recognition, as the representation of a person walking would be deferent to say that of person stood still, or doing some other task.

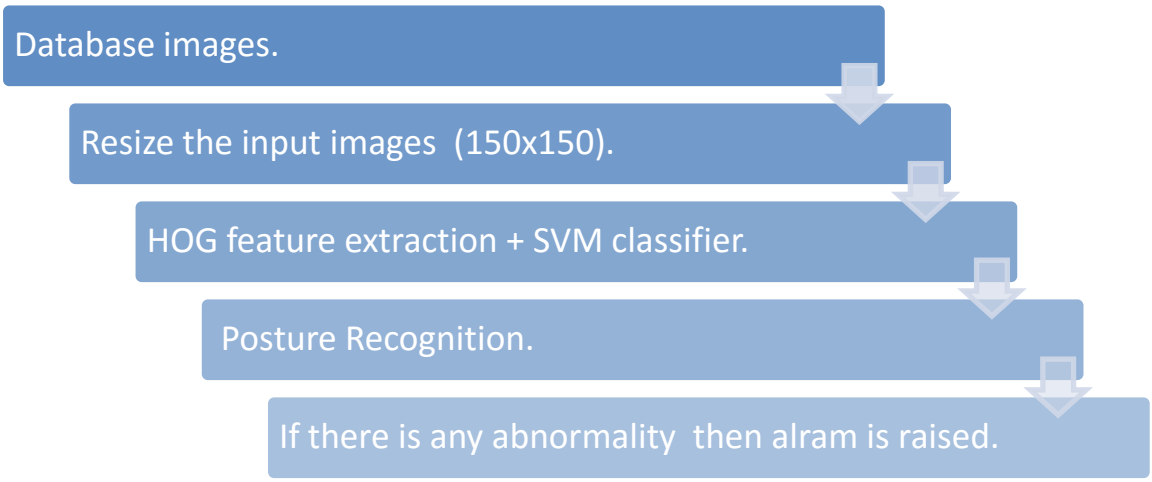


Figure4.1.2: Steps Involved for abnormality recognition

Similarly, the changes in the object shape appearance over time would also draw parallel distinction, however depending on accurate the representation is, somebody walking and somebody stood still could be represented identically. This raises the question of whether a simpler approach, using wrong movement i.e. the area of binary blobs to produce histogram representation would perform just as well as HOG.

Background subtraction is done when we applied it to real-time applications than Background subtraction is used. After the stage of image preprocessing which may include image de-noising, post processing like morphology operations etc. Morphological operations like Dilation, Erosion, Median filter, canny edge detection. Median filter is used to remove the noise from the image. To detect the edges we use canny method. Background subtraction is used for various applications such as abnormal activity detection, gesture recognition, and face recognition methods. Background subtraction is used when we are working in real-time applications. For background subtraction we are using various methods. It includes various steps, first we subtract the original image from the operated image than we convert the image into RGB to grayscale and

then we convert the image into binary image. After that we apply median filter to remove the noise and the we apply thresholding and subtract the background.

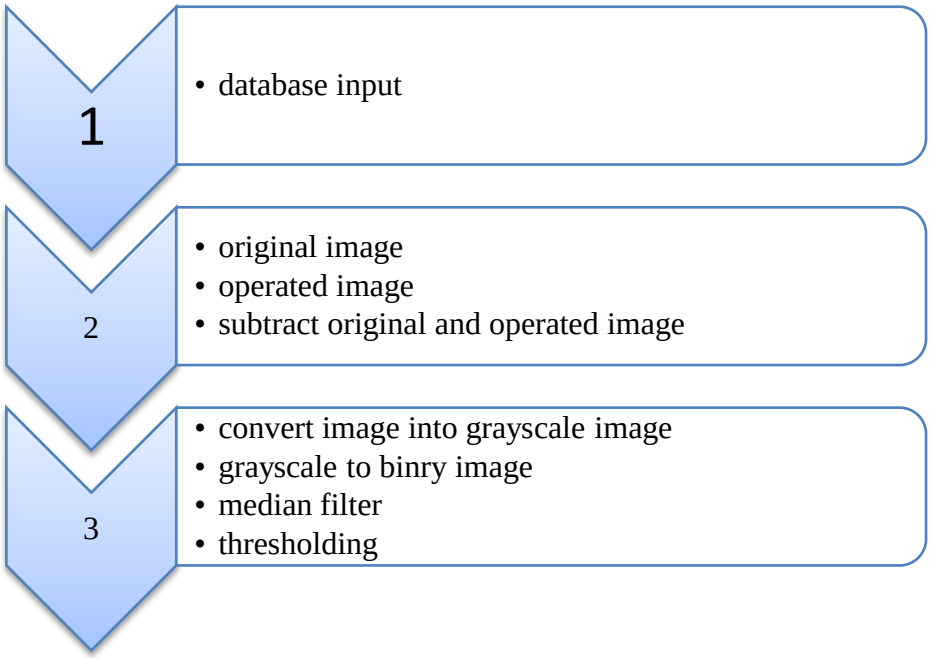


Figure4.1.3: Background Subtraction.

Chapter 5

5.1 Results

In this project, 128 images are taken, thirty eight images of abnormal activity and 90 images of normal activity. Database images are taken as input image, resize the images into 150x150. HOG feature extraction is done.SVM classifier is used for action recognition. If there is any abnormality, then alarm is raised.

Figure5.1.1: Normal Posture

These above images are shown normal postures. These images are taken from the various angles.

Figure5.1.2 Abnormal Posture: (a) person keeping her one hands up,(b) keeping both hands up,(c)T-shape pose,(d)T-shape pose from different angle,(e)another angle of T-shape pose ,(f),(g)different angle of T-shape poses(h)keeping hands half up,(i)standing straight and hands are half up,(j)small degree change and half up ,(k) hands half up (l) standing straight in T-shape pose

The above images are showing abnormal postures. Image are taken from the various angles. these twelve images shows abnormal behavior

Figure5.1.3 Test image: (a)half hands up,(b)different angle of half hands up,(c)T-shape pose ,(d) keeping one hand up,(et-shape pose),(f)different angle of T-shape,(g)different angle of half hand up,(h)keeping hands half up,(i)standing straight,(j)standing straight,(k),standing straight and looking outside the window(l)taking a step

128x46656 double

	1	2	3	4	5	6	7	8
1	0.0576	0.2509	0.3374	0.3374	0.3374	0.0196	0.0054	0.0025
2	0	0	0	0	0	0	0	0
3	0.0112	0.0018	0.1123	0.1575	0.3843	0.1179	0.1767	0.0132
4	0.1722	0.1663	0.1438	0.2068	0.1864	0.1594	0.1140	0.1186
5	0	0	0	0	0.4955	0	0	0
6	0.0047	0.0790	0.1502	0.2791	0.2791	0.0833	0.0221	0.0034
7	0.3282	0.3282	0.1887	0.1231	0.0087	0	0	0.0024
8	0.0755	0.1362	0.1037	0.2754	0.3971	0.1189	0.0633	0.0592
9	0.2869	0.1502	0.0302	0.0012	0.0710	0	0.0119	0.0854
0	3.0219e-04	0.0055	0.0300	0.3836	0.3836	0.0405	0.0025	0
1	0.1775	0.3553	0.0905	0.0022	0.1395	0.0770	0.0244	0.0019
2	0.1722	0.1663	0.1438	0.2068	0.1864	0.1594	0.1140	0.1186
3	0.3007	0.1072	0.0161	0	0	0	0.0017	0.0714

Figure5.1.4: HOG feature extraction

Here, we are taking the test image1, where person is standing straight.

Figure5.1.5: Test image1

1x46656 single

	1	2	3	4	5
1	0.0657	0.0403	0.1473	0.2800	0.2800
2					

Figure5.1.6: Feature extraction of Test image1

Variables - predictedLabels

predictedLabels

1x1 categorical

	1	2	3
1	normal		
2			

Figure5.1.7: Posture Recognition (normal posture)

For Test image2, person is standing toward left and keeping left hand up.

Figure5.1.8: Test image 2

1x46656 single

	1	2	3	4	5
1	0.2549	0.2549	0.0741	0.1081	0.2480
2					

Figure5.1.9: Extracted feature of Test image2

predictedLabels

1x1 categorical

	1	2	3
1	abnormal		
2			
3			

Figure5.1.10: Posture recognition for Test image2 (abnormal activity)

For Test image3, T-shape posture;

Figure5.1.11: Test image3

testfeatures

1x46656 single

	1	2	3	4	5
1	0.3282	0.3282	0.1887	0.1231	0.0087
2					

Figure5.1.12: Extracted feature of Test image3

predictedLabels

1x1 categorical

	1	2	3	4
1	abnormal			
2				
3				
4				
5				

Figure5.1.13: Posture Recognition (abnormal activity)

For Test image4, person is in standing position and joining both hands fingers.

Figure5.1.14: Test image4

testfeatures

1x46656 single

	1	2	3	4	5	6
1	3.4578e-04	0	0.0131	0.1213	0.4069	0.4069
2						
3						

Figure5.1.15: Extracted feature for Test image4.

predictedLabels

1x1 categorical

	1	2	3
1	normal		
2			
3			

Figur5.1.16: Posture Recognition for Test image4 (normal activity).

Figure5.1.17: Test image

testfeatures

1x46656 single

	1	2	3
1	0.0992	0.1791	0.1301
2			
3			
4			
5			

Figure5.1.18: Extracted Feature of Test image5

predictedLabels

1x1 categorical

	1	2	3
1	normal		
2			
3			
4			

Figure5.1.19: Posture Recognition for Test image5

Figure5.1.20: Test imge6

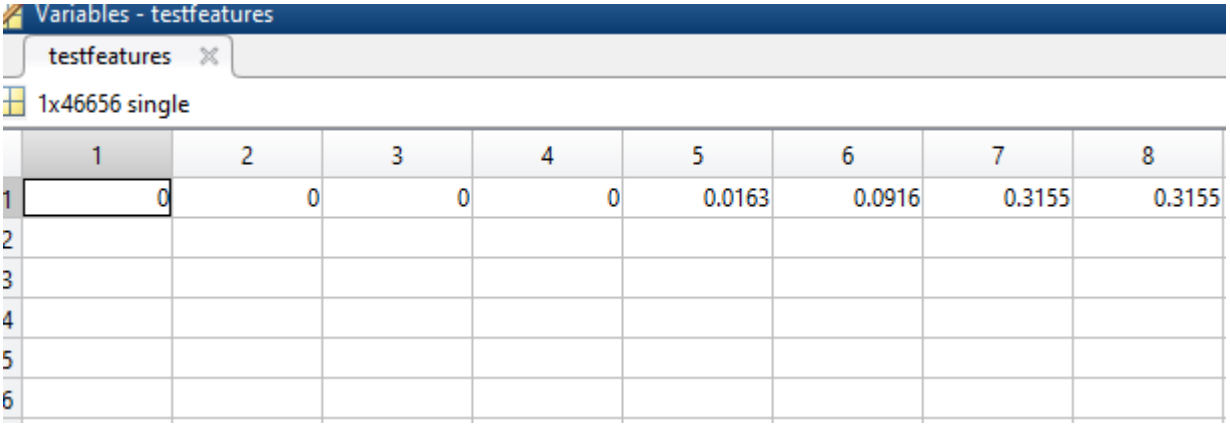


Figure5.1.21: Extracted feature for Test image6

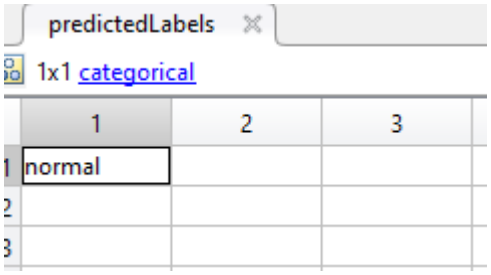


Figure5.1.22: Posture Recognition for Test image6 (normal activity)

5.2 Conclusion

After conducting an offline experimental exploration into the human action recognition problem, it is clear that there are many challenges presented by the realistic environment. The implementation of taking an image and extract actions from the image dataset provides a convenient, time-saving means for building an experimental dataset. It effectively serves the purpose of recognizing the abnormal activity. Histogram of oriented gradient (HOG) is used for feature extraction from the input images and the test images. The experiments conducted in this study consider two action classes; abnormal activity and normal activity. This is a simple binary classification and portrays an intelligent recognition system. Also the actions considered in real-time vary and therefore prove that the system is effective in what is designed for as larger training

set is then available and the experiments gets extended to consider multiple action classes and more postures, to validate the system. Therefore it is clear that an action recognition system would have practical applications in the real world, both offline and in real-time e.g. surveillance system to alarm monitors of suspicious activity.

5.3 Future Prospects

In present thesis we have introduced a method to detect the abnormality recognition in the ATM, by the mean of security bases. At this time we work on static image and we get the result. But for future, we work on real-time bases. We work on real time images (videos).

REFERENCES:

[1] Popoola, Oluwatoyin P., and Kejun Wang. "Video-based abnormal human behaviour recognition—a review." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 42, no. 6 (2012): 865-878.

[2] Gowsikhaa D, Manjunath, Abirami S.”Suspicious Human Activity Detection

[3]Aron F.Bobick and James W.Davis,”The Recognition Of Human Movement Using Temporal Templates”.IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 23, NO., MARCH 2001.

[4]C. Stauffer and W.E.L.Grimson, “Learning Patterns Of Activity Using Real-Time Tracking,”in PAMI,2000.

[5] P Bouthemy and F Meyer, “Region Based Tracking Using Affine Motion Models in Long Image Sequences,”

[6] C. Harris and M.Stephens, “A Combined Corner and Edge Detector,” in AVC,Manchester,1988,pp.147-151.

[7]N.Dalal and B.Triggs, “Histograms of Oriented Gradients for Human Detection,” in CVPR,San Diego,2005,pp.886-893.

[8] N.Dalal , B.Triggs and C.Schmid, “Human Detection using Oriented Histograms of Flow and Apperance,”in ECCV,Graz,2006,pp.428-441.

[9] B.Lucas and T.Kanade, “ An Iterative Image Registeration Technique,” in IJCAI, Vancouver,1981,p.674-679.

[10] B.Horn and B.G.Schunck, “ Determining Optical Flow,” Aritficial Ineteligence, vol.17,pp.185-203,1981.

[11] Chathuramali, KG Manosha, Sameera Ramasinghe, and Ranga Rodrigo. "Abnormal activity recognition using spatio-temporal features." In 7th International Conference on Information and Automation for Sustainability, pp. 1-5. IEEE, 2014.

[12] Santhiya, G., K. Sankaragomathi, S. Selvarani, and A. Niranjil Kumar. "Abnormal Crowd Tracking and motion analysis." In Advanced Communication Control and Computing Technologies (ICACCCT), 2014 International Conference on, pp. 1300-1304. IEEE, 2014.

[13] Tripathi, Vikas, Ankush Mittal, Durgaprasad Gangodkar, and Vishnu Kanth. "Real time security framework for detecting abnormal events at ATM installations." Journal of Real-Time Image Processing (2016): 1-11.

[14] Popoola, Oluwatoyin P., and Kejun Wang. "Video-based abnormal human behaviour recognition—a review." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 42, no. 6 (2012): 865-878.

[15] Rajvi Nar,Alisha Singal and Praveen Kumar, "Abnormal Activity Detection for Bank ATM Surveillance,"Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI), 2016.

[16] Sujith B "Crime Detection and Avoidance in ATM: A NewFramework", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (5) , 2014.

[17] Abdul-Lateef Yussiff, Suet-Peng Yong, Baharum B. Baharudin, "People Detection Enrichment for Abnormal Human Activity Detection**",** Australian Journal of Basic and Applied Sciences, 7(8): 632-640, 2013,ISSN 1991-8178.

[18] Xinyu WU, Yongsheng OU, Huihuan QIAN, and Yangsheng XU, "A Detection System for Human Abnormal Behavior".

[19] Joshua Candamo, Matthew Shreve, "Understanding Transit Scenes: A Survey On Human Behavior-Recognition Algorithms", Vol. 11, No.1, IEEE 2009.

[20] Sugla Vinaayagan Rajenderan, Ka Fei, Thang" Real-Time Detection Of Suspicious Human Movement" SDIWC 2014.

[21] Divya J" Automatic Video Based Surveillance System For Abnormal Behavior Detection" Volume 4 Issue 7, IJSR 2015.

[22] Gowsikhaa D, Manjunath, Abirami S."Suspicious Human Activity Detection From Surveillance Videos" IJIDCS, Vol: 2 No: 2, 2012.

[23] Utkal Sinha, Himanish Shekhar Das and Mayank Shekhar. "Survey on Human Activity Recognition Techniques for Video Surveillance." International Journal of Computer Science and Technology (IJCST), Vol. 6, 2015.

[24] Wren, Christopher Richard, Azarbayejani, Ali, Darrell, Trevor, Pentland, Alex Paul, 1997. Pfinder: Realtime tracking of the human body. Pattern Analysis and Machine Intelligence, *IEEE Transactions on,* 19(7): 780-785.

[25]Jabri, Sumer, Duric, Zoran, Wechsler, Harry, Rosenfeld, Azriel, 2000. Detection and location of people in video images using adaptive fusion of color and edge information. Paper presented at the Pattern Recognition,2000. Proceedings. 15th International Conference on.

[26] Haritaoglu, Ismail, Harwood, David, Davis, Larry S., Real-Time Surveillance of people and their activities, "Pattern Analysis and Machine Intelligence*", IEEE Transactions on,* 22(8): 809-830.

[27]Jabri, Sumer, Duric, Zoran, Wechsler, Harry, Rosenfeld, Azriel, 2000. Detection and location of people in video images using adaptive fusion of color and edge information. Paper presented at the Pattern Recognition, 2000. Proceedings. 15th International Conference on.

[28]Javed, Omar, Shah, Mubarak, 2006. Tracking and object classification for automated surveillanceComputer *Vision—ECCV 2002* (pp. 343-357): Springer.

[29] K. L.K.A.X.Y.Chen, G. Liang, “Abnormal behavior, detection by multi-SVM-based Bayesian network,” in Proc.Int.Conf. Inf. Acquisition, vol. 9, no. 11, July 2007, pp.I–341I–344

[30]F.Lv and R.Nevatia, “Single view human action recognition using key pose matching and vertebra path searching,” in Proc. IEEE Conference. Computer Vision Pattern Recognition, 2007, pp. 1–8.

[31]R.P.Z.Z.Weiyao Lin, Ming-Ting Sun, “Group event detection with a varying number of group members for video surveillance,” in IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY.