Hand Gesture Recognition and Tracking using Image Processing Techniques

A

Dissertation Report

submitted in partial fulfilment of the

requirements for the award of the degree of

MASTER OF TECHNOLOGY

in

ARTIFICIAL INTELLIGENCE AND ARTIFICIAL NEURAL NETWORKS

By

Shardul Singh Chauhan

Under the guidance of

Dr. Venkatadri Marriboyina

Centre for Information Technology

University of Petroleum & Energy Studies

Bidholi, Via Prem Nagar, Dehradun, UK

April – 2016

The innovation driven
E-School

CANDIDATE'S DECLARATION

I hereby certify that the project work entitled **"Hand Gesture Recognition and Tracking using Image Processing Techniques"** in partial fulfilment of the requirements for the award of the Degree of MASTER OF TECHNOLOGY in ARTIFICIAL INTELLIGENCE AND ARTIFICIAL NEURAL NETWORKS and submitted to Centre for Information Technology, University of Petroleum & Energy Studies, Dehradun, is an authentic record of my work carried out during a period from **January, 2016** to **April, 2016** under the supervision of **Dr. Venkatadri Marriboyina, Assistant Professor, CIT, UPES.**

The matter presented in this project has not been submitted by me for the award of any other degree of this or any other University.

(Shardul Singh Chauhan)

Roll No. R102214007

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Date: 11.04.2016

(Dr. Venkatadri Marriboyina)

Project Guide

Dr. Amit Agarwal

Program Head – MTech AI ANN

Centre for Information Technology

University of Petroleum & Energy Studies

Dehradun – 248 007 (Uttarakhand)

ACKNOWLEDGEMENT

I wish to express our deep gratitude to our guide **Dr. Venkatadri Marriboyina**, for all advice, encouragement and constant support he has given me throughout my project work. This work would not have been possible without his support and valuable suggestions.

I am heartily thankful to my course coordinator, **Mr. Vishal Kaushik**, for the precise evaluation of the milestone activities during the project timeline and the qualitative and timely feedback towards the improvement of the project.

I sincerely thank to the respected Program Head of the Department, **Dr. Amit Agarwal**, for his great support in doing my project in **Image Processing** at **CIT**.

I am also grateful to **Dr. Manish Prateek, Associate Dean** and **Dr. Kamal Bansal, Dean CoES, UPES** for giving me the necessary facilities to carry out my project work successfully.

I would like to thank all my **friends** for their help and constructive criticism during my project work. Finally, I have no words to express my sincere gratitude to my **parents** who have shown me this world and for every support they have given me.

Name **Shardul Singh Chauhan**

Roll No. **R102214007**

ABSTRACT

The industry of computers is creating at a quick pace. With this improvement the majority of the fields under computers have progressed in the recent decades. Natural interface with computer using intelligent approaches is the need of Human Computer Interaction (HCI) applications. A presentation of a technique for human computer connection utilizing the client's hand which can be utilized on conventional computer stages. This property thus can be extended in various applications such as in gesture controlled robot or computer, sign language [1] etc. A Real time gesture recognition process utilizes three major steps which includes skin colour detection, extracting convex hull and convexity defect region and convexity defect calculation. The proposed method is able to recognize the number of fingers of a single hand while executing the gestures of hand before a camera device. The aim of the project is to recognize hand gestures and to track the hand movements of an individual in a video.

Keywords: Hand gesture recognition, tracking, computer vision, convex hull, convexity defects, Kalman filter, human computer interaction.

TABLE OF CONTENTS

LIST OF FIGURES

S.No. **Chapter** **Page No.**

1. Introduction

2. System Overview

3. Hand Gesture Recognition and Tracking

4. Implementation

5. Appendix

LIST OF TABLES

1. Introduction

1.1 Human Computer Interface System

Living in cutting edge social orders depends on savvy frameworks and administrations that backing our cooperative and open needs as social creatures. Improvements in programming and equipment innovations as e.g. in discourse innovation, computer semantics, computer vision, and manmade brainpower are persistently driving new applications for work, recreation, and portability. Man-machine collaboration, which is pervasive in human life, is driven by innovation also. It is the portal giving access to capacities and administrations which, because of expanding many-sided quality, undermines to transform into a bottleneck. Since for the client the interface is the item, interface configuration is a key innovation which in like manner is liable to critical national and worldwide exploration endeavours.

Each new device can be seen as an endeavour to make the computer cleverer and making people ready to perform more confused correspondence with the computer. This has been conceivable because of the outcome arranged endeavours made by computer experts for making fruitful human computer interface [1]. As the complexities of all the human needs have transformed into numerous folds and keeps on developing in this way, the requirement for Complex programming capacity and instinct are basic credits of computer software engineers to make due in a focused situation. The computer software engineers have been inconceivably fruitful in facilitating the correspondence in the middle of computers and human. With the rise of each new item in the business sector; it attempts to encourage the diverse nature of all the employments performed. For example, it has helped in encouraging tele working, automated use, better human control over complex work frameworks like autos, planes and observing frameworks. Prior, Computer software engineers were maintaining a strategic distance from such sort of complex projects as the emphasis was more on rate than other modifiable elements. Be that as it may, a move towards an easy to use environment has driven them to return to the centre range [1].

As the computer business takes after Moore's Law since centre 1960s, effective machines are fabricated furnished with more peripherals [3]. Vision based interfaces are practical and at the present minute the computer can "see". Subsequently clients are taking into account wealthier and client friendlier man-machine connection. This can prompt new interfaces that will permit the

arrangement of new summons that are impractical with the present info gadgets. A lot of the reality of the situation will become obvious eventually spared and generally, there has been a surge in excitement for seeing human hand movements. Hand signal affirmation has distinctive applications like computer amusements, contraption control (e.g. crane), and escalated mouse substitution. Today hand motions or any signal utilized as a part of data access can make a framework sufficiently canny to perform the assignment given to it by only a motion and not by utilizing console and mouse. A champion amongst the most composed courses of action of signs has a spot with motion based correspondence. In correspondence by means of motions, each sign has a doled out criticalness (or suggestions).

1.2 Areas of Application of Human Computer Interaction

As indicated by their properties man-machine frameworks might be characterized in dialog frameworks and element frameworks (Fig. 1.1). The previous include little single client machines like cell telephones, individual computerized collaborators and computers, additionally multi-client frameworks like plant control stations or call focuses. These require discrete client activities by, e.g., console or mouse in either stationary or versatile mode.

Agents of the last are area, air, or ocean vehicles, robots, and expert slave frameworks. Taking after the requirements of a maturing society additionally recovery robots, which give development help and reclamation to individuals with inabilities, increment consistently in significance. Every one of these frameworks put the administrator in a control circle and require ceaseless manual control inputs.

Large portions of these require connection with the driver. Soon additionally portable administration robots are normal on the commercial centre as friends or home aides. While these will have worked in insight for to a great extent self-governing operation, they by the by will require human supervision and intermittent mediation.

The portrayal proposed in (Fig. 1.1) rarely happens in unadulterated structure. Numerous frameworks are mixed in restricted or other, requiring discrete dialog activities, manual control, and supervisory control in the meantime. A prototypical illustration is a business carrier, where the cockpit group performs all these activities in succession or, now and again, in parallel.

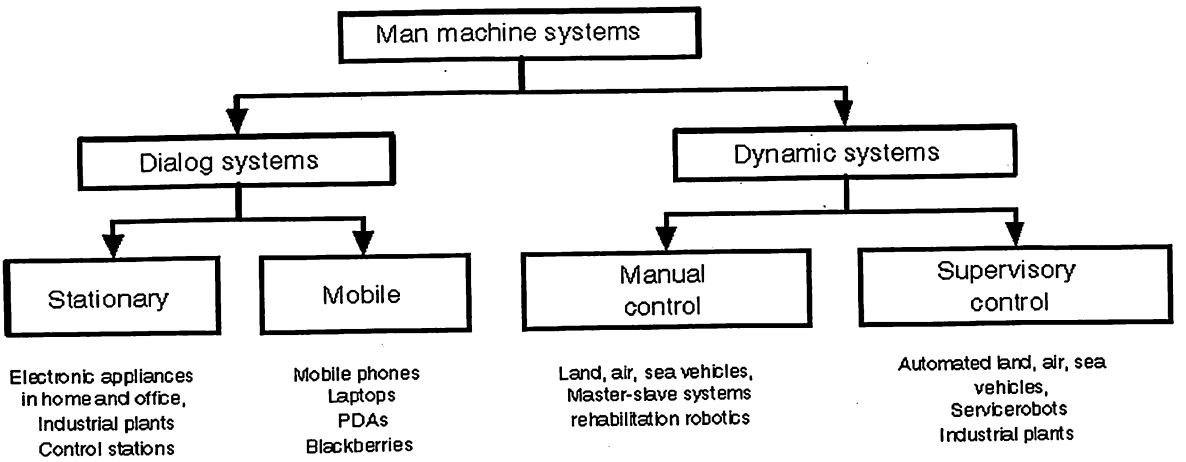


Figure 1-1 Categorization of human computing systems along with their areas of application

As dynamic frameworks are turning out to be more mind boggling, capacities have a tendency to be computerized in parts, which takes the administrator out of the control circle and allocates supervisory control errands to him or recommends participation with machines by assignment sharing. [6] This methodology has a long convention in common and military flight and modern plants. As of now keen robotization installed in this present reality gains fast ground in street activity, where a large number of driver help capacities are a work in progress.

This sort of human machine interface would permit a human client to control remotely through hand signal a wide assortment of gadgets. Distinctive applications have been proposed, for example, contactless control or house apparatuses for welfare change. Another vision based structure is introduced in this strategy, which permits the clients to communicate with computers through hand signal, being the framework versatile to various light conditions and foundations. Its proficiency makes it suitable for constant applications.

1.3 Background

The research work on hand gesture recognition can be divided into three different approaches. The first approach, glove based approach, employs the use of computing devices that is either optical or mechanical. The relative position of the hand is controlled by an extra sensor [2], [8]. This sensor is regularly an attractive or an acoustic sensor connected to the glove. For some

information glove applications, gaze upward table programming toolboxes are given the glove to be utilized for hand stance acknowledgment.

The second classification, vision based examination, depends in transit people see data about their surroundings, yet it is presumably the hardest to execute satisfactorily. A few distinct methodologies have been tried as such. These parameters are then used to perform movement gathering. A hand motion investigation framework taking into account a three-dimensional hand skeleton model with 27 degrees of flexibility was created by Lee and Kunii [7]. They consolidated five noteworthy limitations in light of the human hand kinematics to lessen the model parameter space look. To disentangle the model coordinating, extraordinarily stamped gloves were utilized.

The third classification, examination of drawing motions, for the most part includes the utilization of a stylus as a data gadget [2]. Examination of drawing signals can likewise prompt acknowledgment of composed content. By far most of hand signal acknowledgment work has utilized mechanical detecting, regularly for direct control of a virtual situation and every so often for typical correspondence. Detecting the hand pose mechanically has a scope of issues, be that as it may, including unwavering quality, precision and electromagnetic commotion [5]. Visual detecting can possibly make gestural communication handier, yet conceivably exemplifies the absolute most troublesome issues in machine vision. The hand is a non-unbending protest and much more dreadful self-impediment is extremely normal.

Full ASL acknowledgment frameworks (words, phrases) fuse information gloves. Takashi and Kishino [10] talk about a Data glove-based framework that could perceive 34 of the 46 Japanese motions (client subordinate) utilizing a joint point and hand introduction coding strategy. From their paper, it appears the test client made each of the 46 signals 10 times to give information to rule segment and bunch examination. A different test was made from five emphases of the letters in order by the client, with every signal all around isolated in time. While these frameworks are actually intriguing, they experience the ill effects of an absence of preparing.

Great work has been done in sperling so as to back of machine communication via gestures acknowledgment and Parish [13], who have done watchful studies on the transmission capacity essential for a sign discussion utilizing spatially and transiently sub-inspected pictures. Point light trials (where "lights" are joined to huge areas on the body and simply these focuses are utilized for

acknowledgment), have been done by Poizner [12]. Most frameworks to date study seclude/static motions. In the greater part of the cases those are finger spelling signs.

1.3.1 Object Recognition

Tracking of Large Object

In some intuitive applications [16], the computer needs to track the position or introduction of a hand that is conspicuous in the picture. Pertinent applications may be computer diversions, or intelligent machine control. In such cases, a portrayal of the general properties of the picture, might be satisfactory. Picture minutes, which are quick to figure, give an exceptionally coarse rundown of worldwide midpoints of introduction and position. On the off chance that the hand is on a uniform foundation, this strategy can recognize hand positions and straightforward directing motions [8] [14].

1.3.2 Recognition of shape

Most applications, for example, perceiving specific static hand signal, require a wealthier portrayal of the state of the data object than picture minutes give. In the event that the hand signals fell in a foreordained set, and the camera sees a nearby up of the hand, we might utilize a case based methodology [15], joined with a basic strategy top break down hand signals called introduction histograms. These illustrations based applications include two stages [19]; preparing and running.

1.4 Gestures

It is difficult to settle on a particular valuable meaning of motions because of its wide assortment of utilizations and an announcement can just determine a specific space of motions. Numerous analysts had attempted to characterize signals however their genuine importance is still self-assertive.

Bobick and Wilson [4] have characterized signals as the movement of the body that is planned to correspond with different operators. For a fruitful correspondence, a sender and a recipient must have the same arrangement of data for a specific motion. According to the setting of the undertaking, signal is characterized as an expressive development of body parts which has a specific message, to be imparted decisively between a sender and a recipient. A motion is experimentally arranged into two particular classifications: dynamic and static [1].

A dynamic motion is planned to change over a timeframe though a static motion is seen at the spurt of time. A waving hand implies farewell is a case of element motion and the stop sign is an illustration of static motion. To comprehend a full message, it is important to decipher all the static and element motions over a timeframe. This perplexing procedure is called signal acknowledgment. Motion acknowledgment is the procedure of perceiving and translating a stream constant successive signal from the given arrangement of info information.

1.4.1 Gesture Based Applications

A variety of applications for gesture recognition are broadly classified into two groups which are based upon the different purposes namely a symbolic language and multidirectional control.

3D Design

Computer supported outline (computer offered setup) some assistance with being a HCI which gives a stage to interpretation and control of 3-Dimensional inputs which can be the signs. Controlling 3D inputs with a mouse is a period eating up undertaking as the errand incorporates a confounded methodology of rotting a six-degree opportunity task into no under three progressive two degree assignments. MIT [3] has thought about the 3DRAW development that uses a pen introduced in contraptions to track the pen position and presentation in 3D. A 3space sensor is embedded in a level palette, addressing the plane in which the article or articles rests. A CAD model is moved synchronously with the customer's sign advancements and articles can in this way be turned and made a translation of remembering the final objective to view them from all sides as they are being made and balanced.

Virtual Reality

Virtual the fact of the matter is associated with computer reproduced circumstances that can imitate physical region in spots in this present reality, furthermore in anecdotal universes. Most present virtual reality circumstances are essentially visual experiences, indicated either on a computer screen or through exceptional stereoscopic showcases [6]. There are in like manner a couple of proliferations join additional substantial information, for instance, sound through speakers or headphones. Some best in class, haptic structures now consolidate material information, generally known as force data, in therapeutic and gaming applications.

Sign Languages

Communications via gestures are the rawest and common type of dialects could be gone back to as ahead of schedule as the approach of the human progress, when the principal hypotheses of communications via gestures showed up ever. It has begun even before the rise of talked dialects. From that point forward the gesture based communication has advanced and been received as a fundamental piece of our everyday correspondence process. Presently, gesture based communications are being utilized widely as a part of global sign utilization of not too sharp, in the realm of games, for religious practices furthermore at work places [14]. Motions are one of the principal types of correspondence when a kid figures out how to express its requirement for nourishment, warmth and solace. It improves the accentuation of talked dialect and aides in communicating contemplations and emotions viably.

The acknowledgment of motions speaking to words and sentences as they do in American and Danish communication via gestures [20] without a doubt speaks to the most troublesome acknowledgment issue of those applications said some time recently. A working gesture based communication acknowledgment framework could give a chance to the hard of hearing to correspond with non-marking individuals without the requirement for a mediator. It could be utilized to produce discourse or content making the hard of hearing more free.

The thought is to make computers comprehend human dialect and add to an easy to understand human computer interfaces (HCI). Making a computer comprehend discourse, outward appearances and human motions are a few stages towards it. Signals are the non-verbally traded data. A man can perform endless signals at once. Since human motions are seen through vision, it is a subject of extraordinary enthusiasm for computer vision analysts. The task means to decide human motions by making a HCI in the middle of man and machine.

1.5 Requirement Analysis

The Project requires some essential least assets so it can work appropriately. The Hand Gesture Recognition System has Functional and additionally non-useful prerequisites that must be full dispatched so as to make the application run appropriately. The project needs to full fill the non-useful prerequisites to keep up the nature of its yield and in general. Every one of these necessities have been recorded beneath quickly.

1.5.1 Functional Requirements:

The requirements of the proposed system are summarized below:

- This editor is associated with the framework compiler. Consequently, Visual Studio 2013 or above is required.
- For utilizing the short keys, the working framework backing is required.
- A content tool like Notepad additionally required for altering the info test cases. Information cases in this way are adaptable as goes according to client necessity test cases.

Table 1.1 Essential framework necessities

OS Supported	Windows 7 SP1 (x86 and x64) Windows 8 (x86 and x64) Windows Server 2012 (x64)
Software	Visual Studio 2012 Ultimate for Desktop A software for capturing a video.
Architectures Supported	32-bit (x86) 64-bit (x64)
Other Specific Requirements	1.6 GHz or faster processor 2 GB of RAM 10 GB space of hard disk
Monitor Requirement	15"
Keyboard Requirement	Standard 102 Keys

1.5.2 Non Functional Requirements:

Performance:

The inserted image created ought not contain any twisting. Likewise, the application ought to be secure to factual and correlation examination.

Reliability:

The item ought not crash under any condition, for example, client entering invalid qualities, client attempting to stack unsupported records and so on. It ought to demonstrate suitable message for each client created message.

Portability:

The item will be compact to convey and will keep running in any machine gave it runs a Windows Operating System.

2. Related Work

Research has been constrained to little scale frameworks capable of perceiving an insignificant subset of a full communication through signing. Christopher Lee and Yangsheng Xu [9] added to a glove-based motion acknowledgment framework that could perceive 14 of the letters from the hand letters in order, learn new signals and ready to redesign the model of every motion in the framework in online mode, with a rate of 10Hz. Throughout the years propelled glove gadgets have been outlined, for example, the Sayre Glove, Dexterous Hand Master and PowerGlove [10]. The best industrially accessible glove is by a long shot the VPL DataGlove.

One of the primary expert, computerized presentation apparatuses were the system Harvard Presentation Graphics (Harvard Graphics, 2011), which could be utilized to make slides, diagrams and charts that could then be printed as transparencies. The primary form of Harvard Presentation Graphics was distributed in 1986, at once where computers and projectors were not effectively accessible. The printed transparencies must be put on an overhead projector, which made it bulky to switch forward and backward between slides, and additionally recalling to put the slides on the projector with the right introduction, such that the substance would not be reflected or flipped upside down.

One and only year after the arrival of Harvard Presentation Graphics, Microsoft purchased the rights for a presentation instrument called Presenter, which was created by Bob Gaskins and Dennis Austin, PhD understudies at the University of California-Berkeley. This system was later renamed to PowerPoint, because of copyright issues, and it is this product that is most broadly known and utilized for presentations today (Hewitt, 2008). PowerPoint picks up its quality by the utilization of computers and projectors, which disposes of a ton of the issues that individuals were confronting when utilizing transparencies.

A few papers address the issue of hand and signal acknowledgment, despite the fact that not embedding it in a framework like what Baudel and Beaudouin proposed in their 1993-paper. (Elmezain, Al-Hamadi, Appenrodt, and Michaelis, 2008) made an application that would track a client's hand movement and have the capacity to distinguish the Arabic numbers (0-9) in light of this movement. Elmezain et al. likewise utilizes a profundity division of the followed picture to

have the capacity to all the more effortlessly recognize the territories of significance for following. A strategy that is currently effectively accessible through the equipment of i.e. the Kinect.

Normal for the works said in this part is they all emphasis on how well the innovation functions and how well clients could utilize the created frameworks. Sukthankar et al. inferred that their framework permitted clients to "collaborate with the computer as though it were an individual from the group of onlookers (Sukthankar, Stockton, and Mullin, 2001). Elmezain et al. made an application with high acknowledgment rates for signals and similarly Malik finished up how his hand tracker could track and discover fingertips using a webcam.

Schlömer et al. also, Baudel and Beaudouin-Lafon included clients in their tests and construct their decisions in light of the users" endeavours, yet the tests were just worried with the general population performing the motions and not the ones viewing. Subsequently, it turns out to be less intriguing to examine the utilization of a motion based presentation device from the perspective of the moderator. There has as of now been exploration demonstrating that people are in reality ready to use normal signals and perform important activities with them. Furthermore, fascinating is to take a gander at how the recipients of a motion based presentation will respond and to check whether the utilization of motions changes anything in the group of onlooker's observation and comprehension of the introduced content.

Continuous hand following and 3D motion acknowledgment for intuitive interfaces utilizing HMM [23] The creators have presented a framework which recognizes the hand signals and catches the motions of client having glove with shading. For commotion end this framework utilizes 3D Kalman channels and two shading cameras for 3D recreation.

Model based Segmentation and acknowledgment of element motions in nonstop video stream [21], this paper the creators primarily focused on the division and acknowledgment of ceaseless motions which impact from the spatio-fleeting varieties. The creators utilized two sorts of motions one is two arm developments for form extraction, and another incorporates a solitary hand development hands shape utilized as the component vector. They proposed a Multi scale Gesture Model.

Artificial Neural Network based system for Gesture Recognition is proposed which utilizes Accelerometer Data [26], Accelerometer-Based motion acknowledgment technique is utilized. In second level with no sort of sign preparing for motion acknowledgment Fuzzy automata

calculation has been proposed. In the wake of perceiving the information of the motions, the information was standardized and separated by k-means and Fast Fourier change calculation. Utilizing this Dynamic Bayesian Network, the acknowledgment precision has expanded up to 95%.

Joining different profundity based descriptors for hand motion acknowledgment [22], Based on the profundity data of the picture taken by the profundity cameras the creators have presented a plan known as novel hand signal acknowledgment plan. To legitimately perceive complex signals by utilizing 3-D data they utilized an arrangement of 3-Dimensional components. The proposed hand motion acknowledgment framework comprises of three fundamental steps. The initial step in light of shading and profundity data the hand tests are divided from the foundation.

Entity amputation by patch based filling [11], In this paper another calculation is presented for expelling huge items from advanced pictures. Best first calculation is acquainted with filling the gaps that is cleared out. It is a square based inspecting process, as opposed to an area based procedure. Piece based inspecting procedure is utilized to fill the huge article. The chose expansive article is uprooted and loaded with foundation of the first picture. This strategy takes after a methodology of model based surface combination for filling of the objective locale with a balanced brought together plan.

Another real-time based video system is proposed which uses PixMix [13], High incredibleness constant fit for picture catch and sequence of images. Video deliberation in view of the visual consideration model and internet bunching [18]: Video reflection is utilized to quick search the sequence of images from the huge dataset and to visualize the genuine substance. In light of the Visual consideration and bunching a calculation is proposed. Numerous thoughts are placed in this paper. The main thought is to choose the key edge; a unique substance separation choice standard is utilized.

Multi-sensor foundation subtraction by combining various district based probabilistic classifiers [23], Based on the combination of various locale based classifiers which handle the shading and profundity information which are caught by the RGB cameras, an able foundation subtraction technique is proposed.

The information gathered from these models are utilized by the blend of specialists to expand the closer view discovery exactness. This proposed show essentially concentrates on the closer view and foundation locale based models. An ongoing framework in [31] which tracks the revealed/unmarked hands of a man performing gesture based communication. It extricates the face and hand locales utilizing their skin hues, processes blobs and afterward tracks the area of every hand utilizing a Kalman channel. The framework has been tried for hand following utilizing genuine communication through signing movement by local endorsers. The trial results show that the framework is fit for following hands even while they are covering the face.

Table 2-1 Correlation of various systems for video preparing

APPROACH	**RATE OF EXACTNESS**	**INTENTION**
Video in painting	Intermediate	Stodgy the breach at the forefront.
Chart based subdivision of region.	High	Number of emphases of filling the gaps are diminished.
Best first algorithm, Block based algorithm.	High	Filling the gap for large objects.
Deduction of background method	Intermediate	Filling the gap by region based.
Forefront and circumstantial method, classifiers.	Intermediate	To increase the true detection of the gestures.

Table 2-2 Comparison of different techniques for hand gestures

TECHNIQUE	ACCURACY RATE	AIM
Hand gesture recognition using HMM	98.7%	Spatio temporal variability is reduced.
Multi-scale Gesture Model.	88%-96%	Segmentation of the hand gestures.
Accelerometer-Based gesture recognition	Upto 95%	Normalization, Recognition & filtering the gestures.
Novel hand gesture recognition scheme, SVM classifier.	~95%	3D recognition of hand gestures.
CAMSHIFT algorithm	~93.1%	Recognition, Segmentation and normalization of hand gestures.
Finger Earth Movers distance metric method.	Up to 93.2%	Only fingers of the hand are recognized.

3. System Analysis

3.1 Existing System

Klimis Symeonidis [23] utilized introduction histograms to take a shot at a workstation. It perceives static hand motions, specifically, a subset of American Sign Language (ASL). In this, an example acknowledgment framework is utilizing a change that changes over a picture into an element vector, which will then be contrasted and the element vectors of a preparation set of signals. The last framework is actualized with a Perceptron system.

In [28] a HMD mounted dynamic infrared setup is utilized for fingertip drawing and acknowledgment of items held in the hand. Object acknowledgment is made conceivable by adding a shading camera to the infrared camera on the HMD and utilizing a bar splitter to give the two cameras precisely the same. Thusly the article can be separated from the shading picture utilizing the profundity data from an infrared image.

The Signal Pendant which is exhibited as in [30] is a dynamic infrared camera in an accessory. It is a signal interface essentially intended for home robotization and as a guide for impaired and elderly individuals. Consolidating the dynamic infrared rule with a fisheye lens a great picture of the wearer's hands is acquired. The Gesture acknowledgment is performed with Hidden Markov Models in light of the work done in portable understanding of communication through signing [26].

Hand Gesture Recognition utilizing Computer Vision created by Ray Lockton [22] in which he fabricates a man-machine interface utilizing a camcorder to translate the American one-gave communication through signing letters in order and number signals.

Etsuko Ueda and Yoshio Matsumoto showed a novel methodology a hand-stance estimation that can be used for vision-based human interfaces, in this strategy, the hand locale are removed from various pictures obtained by a multi viewpoint camera structure, and building the "voxel Model" [29].

3.2 Proposed System

The project is divided into two primary modules which are stated below:

- Hand gesture recognition- This module is meant for identifying the hand gestures in the video.
- Tracking- This module is meant for identifying the location of the hand gesture.

First objective of this project is to create a complete system to detect, recognize, interpret and track the hand gesture through computer vision.

Second objective of the project is therefore to provide a low-cost, high speed and colour image acquisition system.

4. Problem Definition

4.1 Problem Statement

A hand signal acknowledgment framework must be made, because of the movements in related fields, especially machine learning, picture get ready and human observation. Similarly, the impact and potential utilization of modified framework have been creating in a broad assortment of uses, including human-computer collaboration, robot control and driver state observation. Then again, to date, energetic acknowledgment of hand motions from images and videos is still a testing errand in view of the inconvenience in unequivocally isolating the accommodating enthusiastic parts. These segments are much of the time addressed in assorted structures, for instance, static, component, point-based geometric or region based elements.

Hand motions acknowledgment and tracking can be performed at various levels of abstraction. In the course of the most recent years, different scientific classifications have been proposed to characterize these levels of deliberation, and diverse names for them have been utilized conversely. The principle point of the undertaken project has been to perceive and track human hand gestures in a video.

4.2 Objectives

The main objectives of the project are stated as under:

- The input to the proposed system is a video which needs to capture the hand movements of an individual.
- To detect hand and categorize hand gestures in the video.
- To track the hand position in the video.
- To provide a low-cost as well as a high speed colour image acquisition system.

5. System Overview

5.1 System Design

Vision based investigation, depends in transit people see data about their surroundings, yet it is presumably the hardest to actualize satisfactorily. A few unique methodologies have been tried in this way.

To catch the picture utilizing a camera then concentrate some element and those components are utilized as data as a part of a characterization calculation for grouping.

5.1.1 Flow Chart

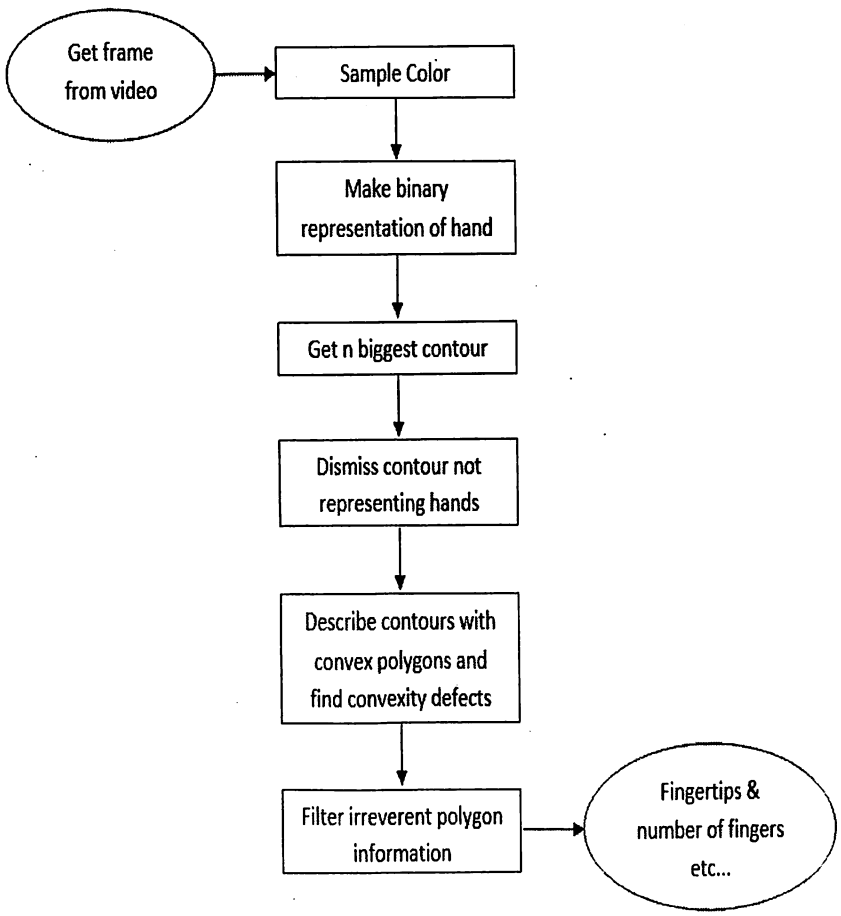


Figure 5-1 Flow chart for hand gesture recognition

Block Diagram

A square outline is a graph of a framework in which the vital parts or capacities are spoken to by pieces associated by lines that demonstrate the connections of the squares. They are intensely utilized as a part of building in equipment outline, electronic configuration, programming plan, and process stream graphs.

Piece outlines are normally utilized for larger amount, less nitty gritty depictions that are proposed to illuminate general ideas without sympathy toward the points of interest of usage. Balance this with the schematic outlines and design graphs utilized as a part of electrical building, which demonstrate the usage points of interest of electrical segments and physical development.

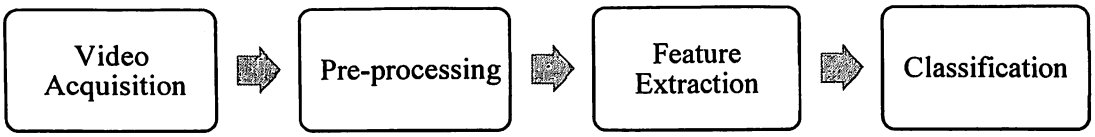


Figure 5-2 Block Diagram of hand gesture recognition system

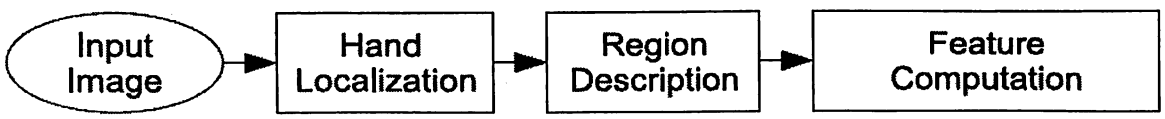


Figure 5-3 Block Diagram for Feature extraction

5.1.2 Data Flow Diagram:

A data flow diagram (DFD) is a graphical representation of the "stream" of data through an information system, showing its methodology perspectives. A DFD is used as a preliminary step to make a survey of the structure, which can later be clarified. DFDs can similarly be used for the view of data get ready.

DFD Level 0:

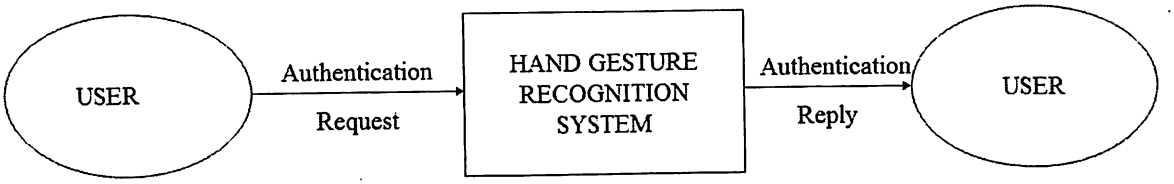


Figure 5-4 Level-0 DFD Diagram

DFD Level 1:

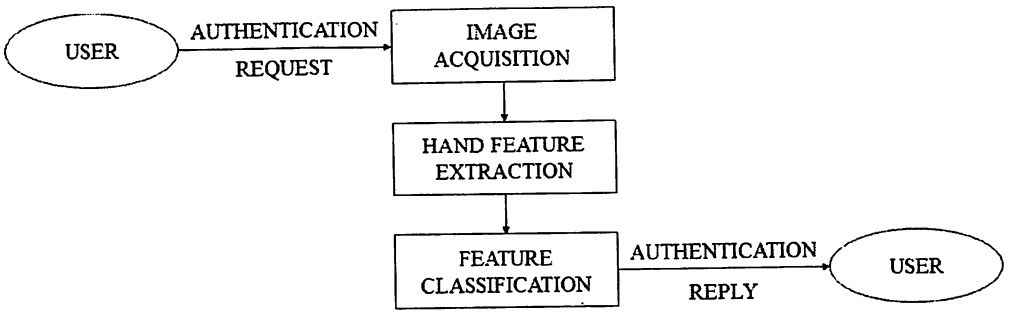


Figure 5-5 DFD Level-1

5.1.3 Use Case Diagram

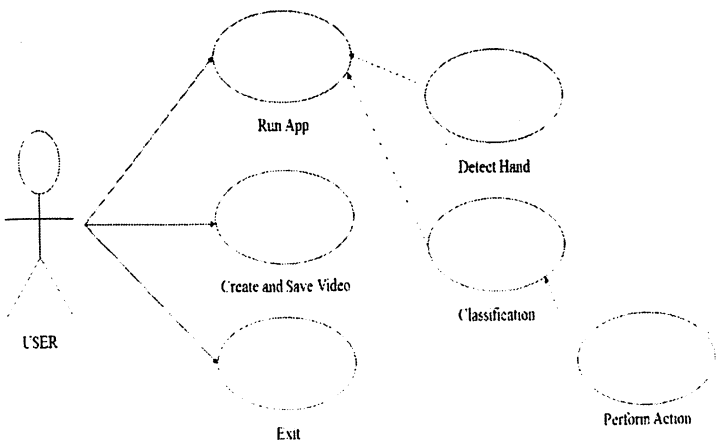


Figure 5-6 Use Case Diagram for the proposed system

5.1.4 Activity Diagram

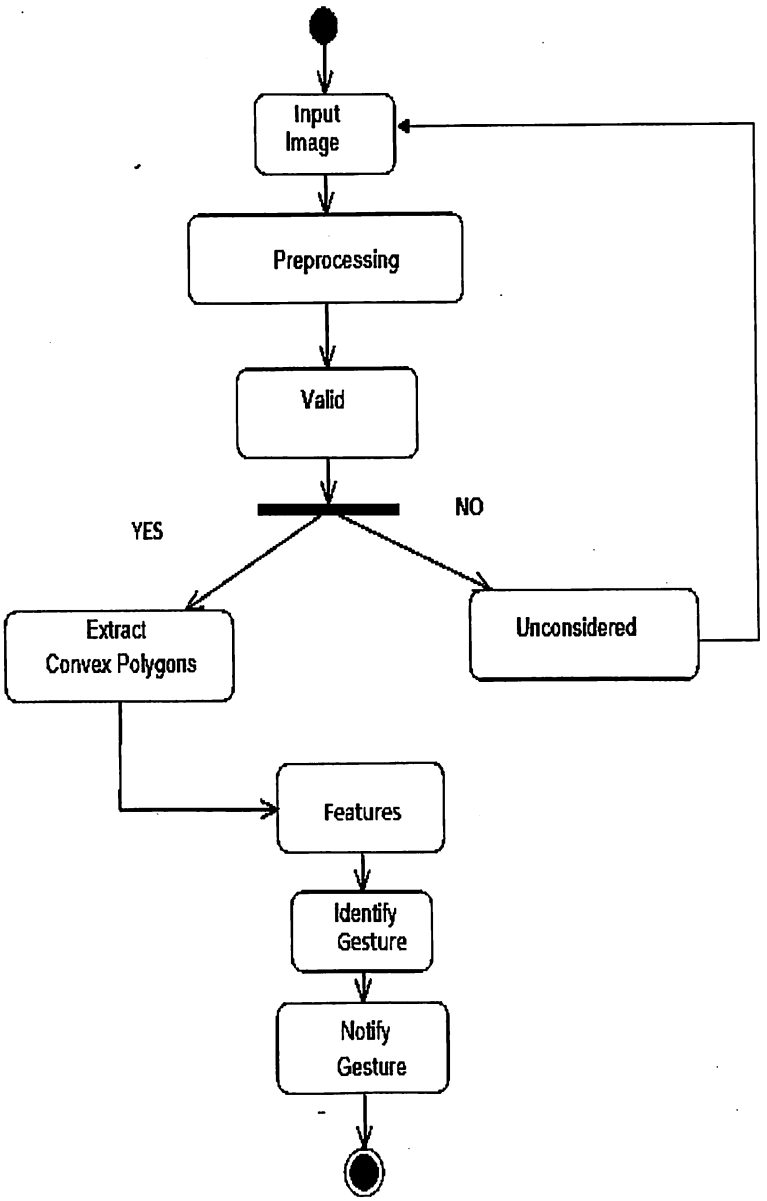


Figure 5-7 Activity diagram for the system

5.1.5 Collaboration Diagram

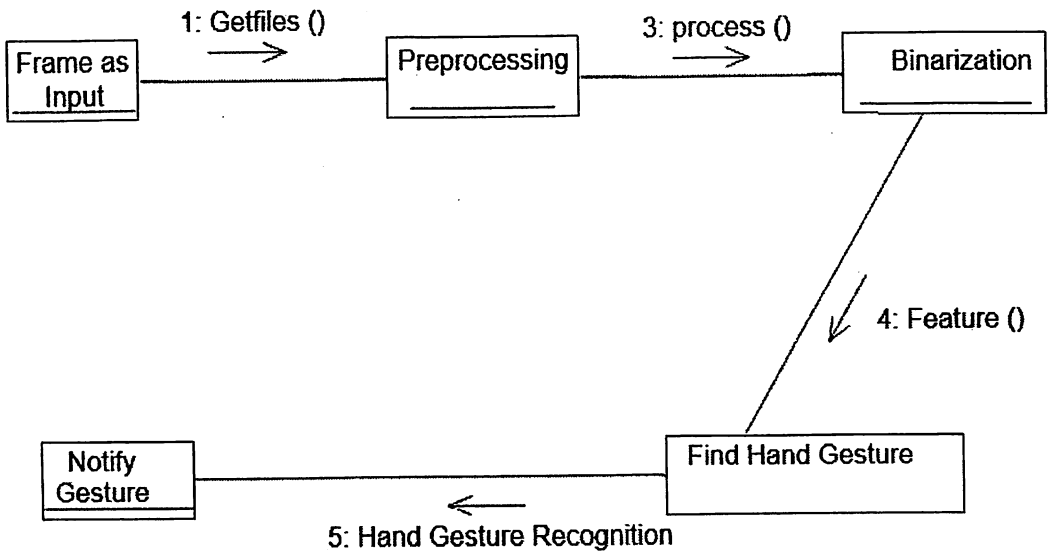


Figure 5-8 Collaboration Diagram for the proposed system

6. Hand Gesture Recognition and Tracking

6.1 Acquisition and Input Data

Before investigating calculations, it is essential to look at those parts of the arranged signal acknowledgment framework that straightforwardly influence the communication with the client. Three vital properties can be characterized that extraordinarily impact the ensuing configuration process:

- The vocabulary, i.e. the arrangement of motions that are to be perceived. The vocabulary size is typically indistinguishable with the quantity of conceivable framework reactions (unless various signals trigger the same reaction).
- The recording conditions under which the framework works.
- To disentangle framework outline, confinements are ordinarily forced on picture foundation, lighting, the objective's base and most extreme separation to the camera, and so on.

6.2 Representation

The depiction and usage of picture handling calculations requires a suitable numerical representation of different sorts of pictures. A typical representation of a rectangular picture with a determination of M lines and N segments is a discrete capacity

I (x, y) with x ∈ {0, 1. . . N 1}; y ∈ {0, 1. . . M 1}

The 2-tuple (x, y) signifies pixel organizes with the birthplace (0, 0) in the upper left corner. The estimation of I (x, y) depicts a property of the relating pixel. This property might be the pixel's shading, its splendour, the likelihood of it speaking to human skin, and so on.

Shading is depicted as a n-tuple of scalar qualities, contingent upon the picked shading model. The most well-known shading model in computer design is RGB, which utilizes a 3-tuple (r, g, b) indicating a shading's red, green, and blue segments:

$$\mathbf{I}(x,y) = \begin{pmatrix} r(x,y) \\ g(x,y) \\ b(x,y) \end{pmatrix}$$

Other as often as possible utilized shading models are HSI (tint, immersion, power) or YCbCr (luminance and chrominance). For shading pictures, I (x, y) is consequently a vector capacity. For some different properties, for example, brilliance and probabilities, I (x, y) is a scalar capacity and can be pictured as a dim worth picture (likewise called force picture). Scalar whole numbers can be utilized to encode a pixel's grouping. For signal acknowledgment a critical arrangement is the segregation between closer view (target) and foundation. This yields a double esteemed picture, generally called a veil,

$I_{mask}(x, y) \in \{0, 1\}$ *Binary values are usually visualized as black and white.*

6.3 Extraction of features

The move from low-level picture information to some more elevated amount depiction thereof, spoke to as a vector of scalar qualities, is called highlight extraction. In this procedure, insignificant data (foundation) is disposed of, while important data (forefront or target) is segregated. Most example acknowledgment frameworks perform this stride, since handling the complete picture is computationally excessively requesting and presents an unsatisfactorily high measure of clamour (pictures regularly contain essentially more foundation pixels than closer view pixels).

Utilizing two-dimensional pictures of the unmarked hand, it is at present impractical to make a three-dimensional model of a deformable item as unpredictable and adaptable as the hand continuously. Since the hand has 27 degrees of flexibility (DOF), this would require a measure of data that can't be removed with adequate exactness and pace. Model-based components, for example, the bowing of individual fingers are hence not accessible in video-based without marker continuous motion acknowledgment. Rather, appearance-based components that portray the two-dimensional perspective of the hand are misused.

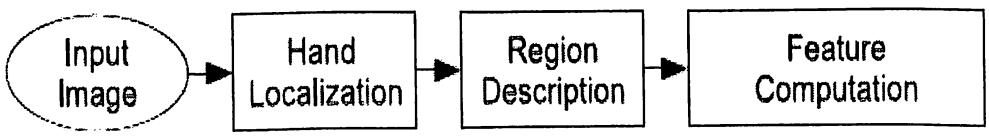


Figure 6-1 Feature extraction of frame

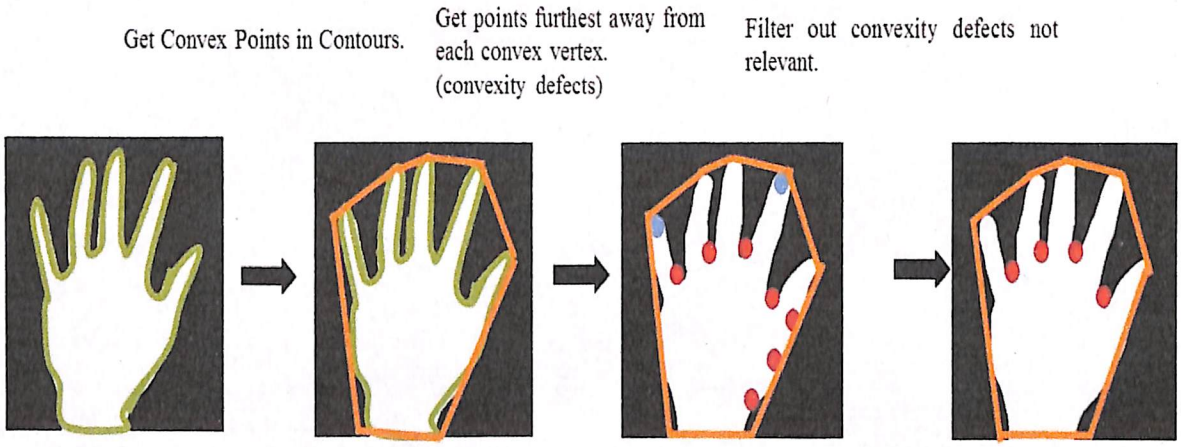


Figure 6-2 When the binary representation is generated the hand is processed in the described way

6.4 Localization of hands

The recognizable proof of closer view or target locales constitutes an understanding of the picture in light of information which is normally particular to the application situation. This learning can be encoded expressly or certainly. Known properties of the objective article, for example, shape, size, or shading, can be misused. In motion acknowledgment, shading is the most every now and again utilized element for hand restriction since shape and size of the hand's projection in the two dimensional picture plane differ extraordinarily. It is likewise the main component unequivocally put away in the picture.

Utilizing the shading ascribe to restrict an item in the picture requires a meaning of the article's shading or hues. In the RGB shading model (and most others), even protests that one would call uncoloured for the most part involve a scope of numerical qualities. This reach can be portrayed factually utilizing a three-dimensional discrete histogram h object (r, g, b), with the measurements relating to the red, green, and blue segments.

$$\sum r \sum g \sum b \; h \; \text{object}\,(r, g, b) = n \; \text{object}$$

6.4.1 Inversion of Colour of Skin

Skin-shading Conversion method is considered as a successful device for face identification since it is invariant to changes in size, introduction and impediment. I propose to utilize the YCbCr shading space for two reasons:

- By utilizing YCbCr shading space, we can wipe out however much as could reasonably be expected the variety of luminance part brought about by the lighting condition.
- The YCbCr area is widely utilized as a part of computerized video coding applications. YCbCr is a shading space that isolates the luminance from the shading data. Luminance is encoded in Y component, blueness in the Cb component and redness in Cr component. It is anything but difficult to change over from RGB to YCbCr.

Any RGB digital image can be converted into YCrCb colour space using following equations:

$$Y = 0.299\,R + 0.587\,G + 0.114\,B$$

$$Cb = -0.169\,R - 0.331\,G + 0.500\,B$$

$$Cr = 0.500\,R - 0.419\,G - 0.081\,B$$

6.5 Region Description

On the premise of Iobj, cover the source picture I can be portioned into locales. An area R is a bordering set of pixels' p for which Iobj, veil has the same worth. The idea of contiguity requires a meaning of pixel nearness. As portrayed in Fig. 6.3, nearness might be founded on either a 4-neighborhood or an 8-neighborhood. In this area the 8-neighborhood will be utilized. By and large, areas might contain different locales and/or openings, however this should not be considered here on the grounds that it is of minor significance in most motion acknowledgment applications.

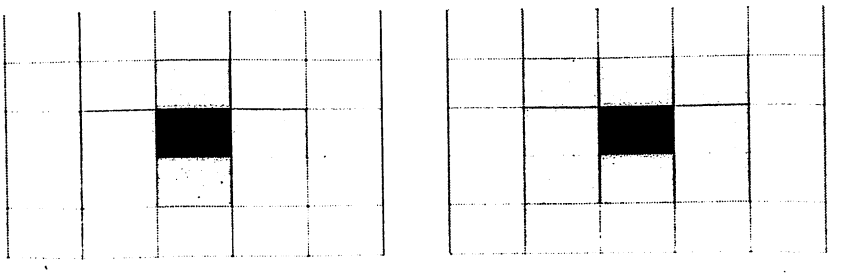
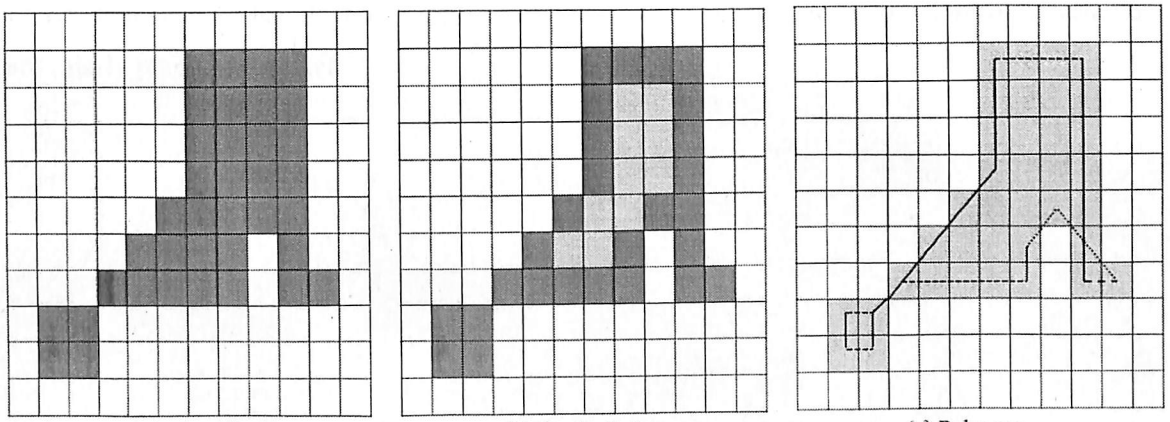


Figure 6-3 Adjacent pixels (gray) in the 4-neighborhood (a) 8-neighborhood (b)

Under lab recording conditions, I obj, veil will contain precisely one target district. In any case, in numerous certifiable applications, other skin hued items might be unmistakable too, so that Iobj, cover ordinarily contains different target districts. Unless propelled thinking procedures are utilized, the component extraction stage is in charge of recognizing the area that speaks to the client's hand, conceivably among applicants that are large in numbers. This is usually done on premise of the area's geometric elements.

In a counter clockwise traversal of the item's fringe, each outskirt point has an antecedent and a successor inside of its 8-neighborhood. A proficient information structure for the representation of a locale is a sorted rundown of its outskirt focuses. This can be deciphered as a shut polygon whose vertices are the focuses of the fringe focuses (Fig. 6.4c). In the accompanying, the article's outskirt is characterized to be this polygon. This definition has the upsides of being sub-pixel exact and encouraging productive calculation of different shape-based geometric elements.



(a) Region Pixels (b) Border Points (c) Polygons

Figure 6-4 Description of an image region by the set of all of its pixels (a), a list of its border pixels (b), and a closed polygon (c)

6.6 Classification of features

The assignment of highlight arrangement happens in all example acknowledgment frameworks and has been subject of significant exploration exertion. An expansive number of calculations is accessible to manufacture classifiers for different necessities. The classifiers considered in this segment work in two stages: In the preparation stage the classifier "takes in" the vocabulary from an adequately extensive number of agent illustrations (the preparation tests). This "information"

is then connected in the accompanying characterization stage. Classifiers that keep on adapting even in the characterization stage, e.g. to consequently adjust to the present client, might not be considered here.

6.7 Kalman Filter for tracking

The Kalman Filter, generally called straight quadratic estimation (LQE), is a calculation that uses a movement of estimations saw after some time, containing confusion (sporadic assortments) and diverse oversights, and conveys evaluations of dark variables that tend to be more correct than those in perspective of a lone estimation alone. More formally, the Kalman channel works recursively on surges of rambunctious data to make a quantifiably perfect appraisal of the fundamental structure state. The channel is named for Rudolf (Rudy) E. Kálmán, one of the vital architects of its theory.

The Kalman channel has different applications, e.g., for bearing, course and control of vehicles, particularly plane and rocket.

7. Implementation

Interpretation of graphical data comprises evaluation of frames taken under diverse settings and at dissimilar moments in time. It can be supposed to be the utmost elementary stage in completing a productive framework and in giving the client, certainty that the new framework will work and be viable.

Hand Gesture Recognition system can be divided into following modules:

- Pre-processing
- Feature Extraction of the processed Image
- Real time classification

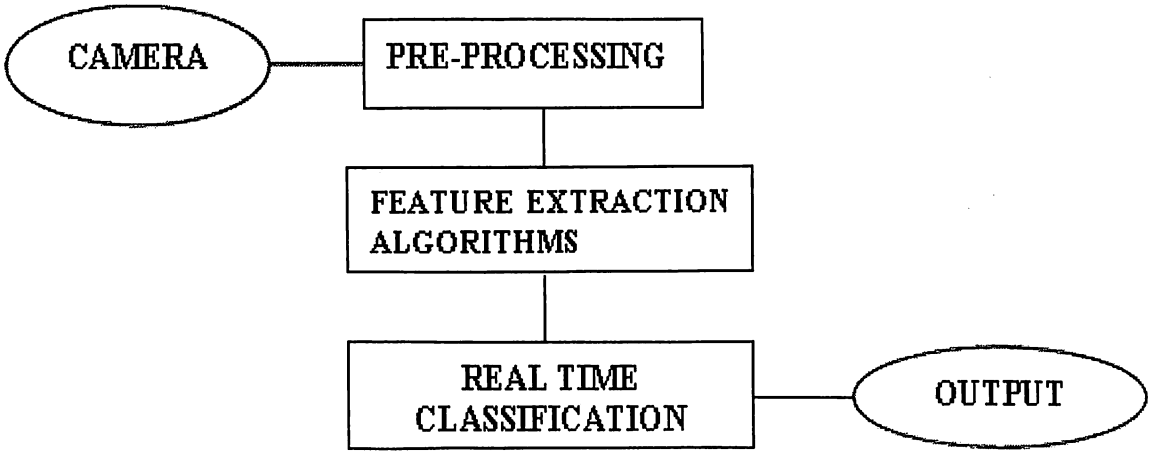


Figure 7-1 System Implementation

7.1 Pre-processing

Like many other pattern recognition tasks, pre-processing is necessary for enhancing robustness and recognition accuracy.

The pre-processing prepares the image sequence for the recognition, so before calculating the diagonal sum and other algorithms, a pre-processing step is performed to get the appropriate image, which is required for real time classification. So it consists of some steps. The net effect of this processing is to extract the hand only from the given input because once the hand is detected

from the given input it can be recognized easily. So pre-processing step mainly consists of the following tasks:

- Skin Modelling
- Removal of Background
- Conversion of RGB to Binary
- Hand Detection

7.1.1 Skin Modelling

There are numerous methods used for skin detection such as RGB, YCbCr and HSV.

RGB

RGB is a 3D colour space pixel where each pixel has combination of three colours Red, Green and Blue at specific location. This technique is widely used in image processing for identifying skin region.

YCbCr (Luminance Chrominance)

This colour space is used in digital video colour information, to represent two colour Cb and Cr. Cb is difference between Blue and Cr is difference between Red component references of value. This is basically RGB transformation to YCbCr for separation of luminance and chrominance for colour modelling.

HSV (Hue, Saturation and Value)

In HSV, Hue detect dominant colour and saturation defines colourfulness whilst Value measure intensity or brightness. This is well enough to choose single colour but it ignores complexity of colour appearance. It trades off computation speed mean computationally expensive and perceptual relevance.

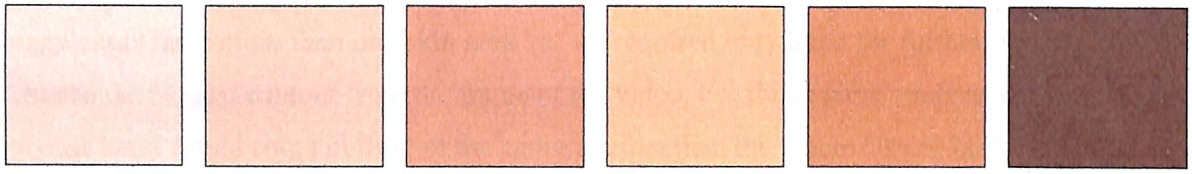
My approach for this project is to work with RGB to binarization techniques to explicitly defined skin region.

Skin Detection

The skin colour detection is one of important goal in hand gesture recognition. Skin colour detection decision rules which we have to build that will discriminate between skin portion and non-skin portion pixels. This is accomplished usually by metric introduction, which measure distance of the pixel colour. This metric type is known as skin modelling.

Explicitly Defined Skin Region

Following are some common ethnic skin groups and there RGB colour space:



European	Middle Eastern	Eastern	Asian	Lt. Black	Dk. Black
R=245 G= 218 B=204	R=237 G= 191 B=166	R=211 G= 141 B=111	R=233 G= 183 B=138	R=197 G= 132 B=92	R=96 G= 59 B=43

Figure 7-2 Different Ethnic Group Skin Patches

To assemble a skin classifier is to characterize unequivocally through various standards the limits of skin shading group in some shading space. The upside of this strategy is the straightforwardness of skin identification decides that prompts the development of exceptionally quick classifier. For e.g.

(Red, Green, Blue) is categorized as skin if and only if:

Red>67 and Green>34 and Blue>35 and

Max {Red, Green, Blue}-min {Red, Green, Blue} >23 and

|Red-Green| > 23 and Red>Green and Red>Blue

In this classifier threshold defined to maximize the chance for the skin region for each colour. If we see fig 7.2, that Red colour in every skin sample is greater than 95, Green is greater than 40

and Blue is greater than 20 in. So threshold can make this classifier easily detect almost all kind of skin.

7.1.2 Removal of Background

Conversion of RGB to Binary

All algorithm accepts an input in RGB form and then convert it into binary format in order to provide ease in recognizing any gesture and also retaining the luminance factor in an image.

Hand Detection

Image could have more than one skin area but we required only hand for further process. For this I choose the biggest contour from the frame of the video. For this, I have made an assumption that only the hand would come in front of the camera rather than the whole human body. If in case that happens then the system will find the biggest contour from the frame and computes likewise. Thus chances of recognizing the gesture would be very difficult. As the system is not ready for the real time application, however further development can be done in this to implement it in real time.

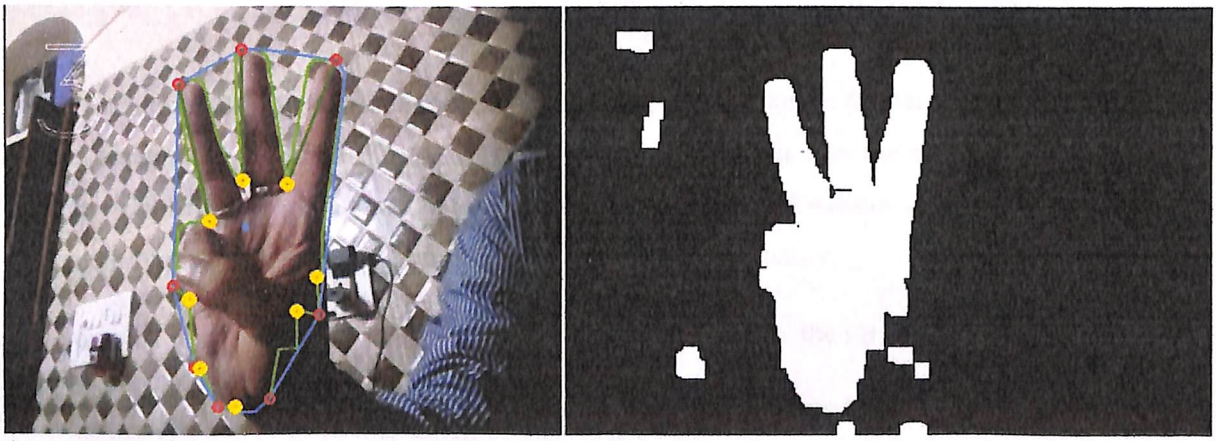


Figure 7-3 Skin Region is represented as the Binary Image after morphological operations

7.2 Feature Extraction

7.2.1 Morphological Operations

The progressions given by the foundation subtraction is light and is high in commotion. To stifle the clamour, we utilize a mix of disintegration and expansion. These are morphological changes.

Expansion operations comprises of convoluting a picture with some portion, which can have any shape or size, as a rule a square or circle. The portion has a characterized grapple point, for the most part being the focal point of the piece. As the bit is looked over the picture, we process the maximal pixel esteem covered by and supplant the picture pixel in the grapple point position with that maximal worth. As you can reason, this expanding operation causes splendid districts inside of a picture to develop.

We can think about a parallel picture I (u, v) as the arrangement of all pixel areas in the closer view:

$$Q_1 = \{(u, v) \mid I (u, v) = 1\}$$

To disentangle documentation, we'll utilize a solitary variable for a direction pair, p = (u, v). Along these lines,

$$Q_1 = \{p \mid I (p) = 1\}$$

A widening of a picture I by the structure component H is given by the set operation:

$$I \ddot{o} H = \{(p + q) \mid p \varepsilon I, q \varepsilon H\}$$

The foundation (brilliant) widens around the dark areas of the letter. At that point we now perform disintegration on it. What this does is to register a nearby least over the zone of the part. As the piece is looked over the picture, we register the negligible pixel esteem covered by and supplant the picture pixel under the stay point with that insignificant quality.

An erosion of an image I by the structure element H is given by the set operation.

$$I \bar{o} H = \{p \varepsilon Z^2 \mid (p + q) \varepsilon I, \text{ for every } q \varepsilon H\}$$

7.2.2 Contour Extraction & Equalization

At in the first place, we characterize such an item form. The shape is a limit of article, a populace of focuses (pixels), and isolating item from a foundation.

In frameworks of computer vision, a few arrangements of coding of a shape are utilized - the code of Freeman, two-dimensional coding, polygonal coding are generally known.

Rather, in a CA the form is encoded by the succession comprising of complex numbers. On a shape, the point which is called as beginning stage is settled. At that point, the shape is filtered (is permissible - clockwise), and every vector of balance is noted by a perplexing number a+ib. Where a - point balance on x pivot, and b - balance on y hub. Counterbalanced is noted concerning the past point.

Every vector of a form we will name basic vector (EV). Furthermore, grouping of complex-esteemed numbers - vector-shape (VC).

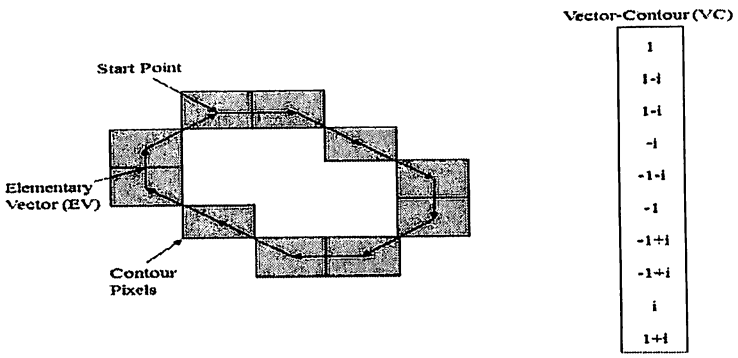


Figure 7-4 : Evaluation of contour using freeman chain algorithm

Thus, vector-contour Γ of length k can be designated as:

$$\Gamma = (\lambda 0, \lambda 1, \lambda_2 \ldots\ldots \lambda k\text{-}2, \lambda k\text{-}1)$$

This gives us a list of set of points, each set representing a contour. We can filter out small contours as they will be noise.

In genuine, image forms have discretionary length. In this way, to search and looking at of forms, every one of them ought to be directed to uniform length. This procedure is called evening out.

- At to start with, we alter length of a vector shape which we will use in our arrangement of an acknowledgment. We assign it k.
- Then, for every underlying form "A" we make vector-shape "N" along 'k'. Assist most likely two variations - or the underlying form has more prominent number of a basic vector than 'k', or littler number than 'k'.

- If an underlying shape more fundamental it is sorted out all by its basic vector, and we consider components "N" as the aggregate of every rudimentary vector.

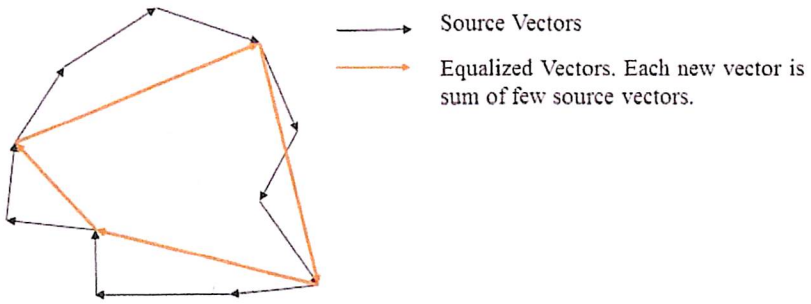


Figure 7-5 Convex Hull & Convexity Defects

Presently given the arrangement of focuses for the shape, we locate the littlest zone arched body that covers the forms. The perception here is that the arched frame focuses are well on the way to be on the fingers as they are the furthest points and thus this can be utilized to recognize number of fingers. Be that as it may, subsequent to our whole arm arrives, there will be different purposes of convexity as well. So we locate the curved imperfections i.e., between every arm of the frame, we attempt to locate the most profound purpose of deviation on the form.

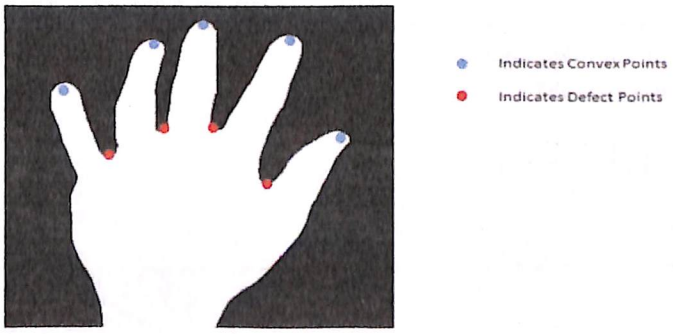


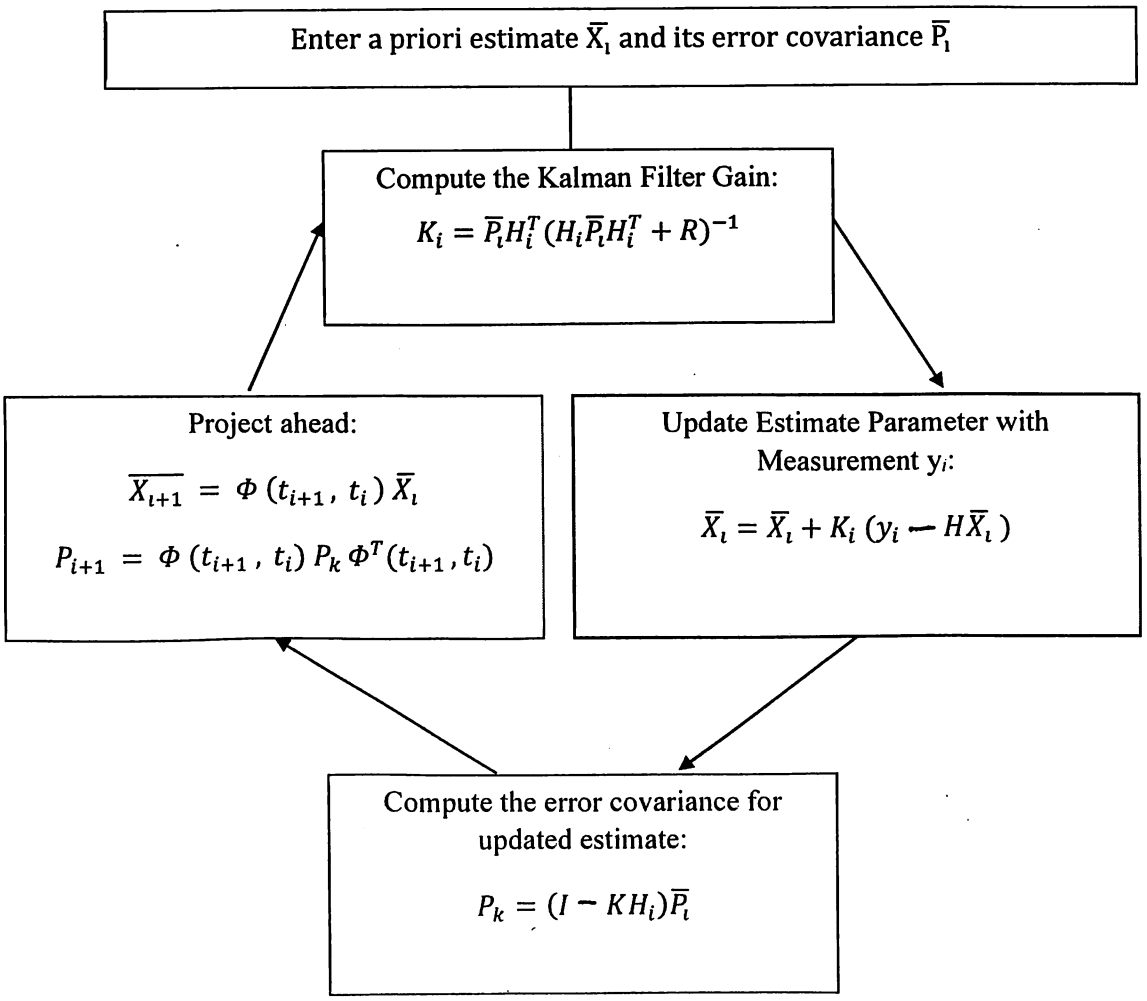
Figure 7-6 Contours

7.3 Finger Detection

The gesture for finding the number of fingers can be calculated by finding the distance between convex point and defect point of the contour. The implementation is described below:

7.4 Tracking

For the purpose of tracking Kalman filter is used. The calculation works in a two-stage process. In the desire step, the Kalman channel produces evaluations of the present state variables, nearby their vulnerabilities. Once the after-effect of the accompanying estimation (basically demolished with some measure of botch, including self-assertive tumult) is viewed, these appraisals are updated using a weighted typical, with more weight being given to assesses with higher conviction. The way of the calculation is recursive. It can continue running dynamically, using only the present information estimations and the previously found out state and its helplessness system; no additional past information is required. The Kalman Filter Algorithm is described mathematically as follows:



7.5 C#

In June 2000, Microsoft declared both the .NET stage and another programming dialect called C#. C# is a basic, cutting edge, object situated, and sort safe programming dialect got from C and C++. C# (declared" C sharp") is solidly planted in the C and C++ family tree of dialects, and will instantly be well known to C and C++ software engineers. C# means to consolidate the high profitability of Visual Basic and the crude force of C++. C# is a specifically question situated dialect intended to give the ideal mix of straightforwardness, expressiveness, and execution. The .NET stage is focused on a Common Language Runtime (like a JVM) and an arrangement of libraries which can be abused by a wide assortment of dialects which can cooperate by all aggregating to a middle of the road dialect (IL). C# and .NET are somewhat harmonious: a few elements of C# arrive to function admirably with .NET, and a few components of .NET arrive to function admirably with C# (however .NET expects to function admirably with numerous dialects).

7.5.1 C# and its features

- C# is a basic, cutting edge, object situated dialect got from C++ and Java.
- It means to consolidate the high profitability of Visual Basic and the crude force of C++.
- It is a piece of MS Visual Studio version 7.0.
- Visual studio bolsters Vb, C++, VC++, Vbscript, and JScript. These dialects give access to the Microsoft .NET stage.
- DOT NET incorporates a Common Execution motor and a rich class library.

8. Limitations & Future Enhancements

8.1 Limitations

- The application has been tested for Windows platform only.
- The application can be enhanced using a different and more efficient algorithm. Although the algorithm is efficient but it can be further optimized to produce desired results.
- The input video should have the hand region occupying major area in the overall frame.
- A less number of hand gestures are being identified by the proposed system.
- The overall system is not suitable to implement in the real time environment due to the limitation of the hand region.
- Due to skin colour based extraction method, if any object is having the similar RGB values the system couldn't produce the satisfactory results.

8.2 Future Enhancements

- The application can be made to work on other platforms also by downloading the libraries of that platform and implementing them in coding.
- A more optimized algorithm can be used in the system such that the accuracy of the system can be enhanced.
- A dynamic system can be modelled so as to work in real time environment.
- The system can be made to recognize sign language or hand gesture based operations for real time environment.

9. Conclusion

In past years a considerable amount of research has been conducted in gesture recognition. The aim of this project is to develop an offline Gesture recognition system. This project is to device an application that can able to detect hands from a video frame, track them in real time and perform gesture recognition. In this project, we proposed a new approach to recognize few of the hand signs using image processing techniques. It is resolved that shape is imperative element and can be utilized for segregation between two signal. The handling ventures to characterize a motion included signal securing, division, morphological separating, form representation and arrangement utilizing distinctive method. In this, system takes video as an input after which it extracts the region of interest using colour spaces (YCrCb). The binary mask thus produced is further processed so as to remove the noise by applying image processing techniques. The convex hull and convexity defect points are then computed so as to calculate the number of fingers from the binary mask. There is a lot of scope for the project to further enhance, for e.g. by improving the colour spaces detection and trying out the project to recognize the sign language.

References

[1] Prof. Dr.-Ing. Karl-Friedrich Kraiss, "Advanced Man-Machine Interaction_ Fundamentals and Implementation", Springer.

[2] Rafael C. Gonzale, Richard E. Woods, "Digital Image Processing", 3rd Edition, Pearson Education.

[3] Al Bovik, "The Essential Guide to Image Processing", 2nd Edition, Elsevier.

[4] S. Annadurai, R. Shanmugalakshmi, "Fundamental of Digital Image Processing", 1st Edition, Pearson Education.

[5] Christopher Lee, Yangsheng Xu, "Online, interactive learning of gestures for human robot interfaces", Carnegie Mellon University, The Robotics Institute, Pittsburgh, Pennsylvania, USA, 1996.

[6] Richard Watson, "Gesture recognition techniques", Technical report, Trinity College, Department of Computer Science, Dublin, July, Technical Report No. TCD-CS-93-11, 1993.

[7] Lee and Kunii, Ray Lockton," Hand Gesture Recognition using computer Vision",4th year project report, Ballilol College, Department of Engineering Science, Oxford University,2000.

[8] Ms. Rashmi D. Kyatanavar, Prof. P. R. Futane, "Comparative Study of Sign Language Recognition Systems", Department of Computer Engineering, Sinhgad College of Engineering, Pune, India International Journal of Scientific and Research Publications, Volume 2, Issue 6, June 2012 ISSN 2250.

[9] International Multi Conference of Engineers and Computer Scientists 2009, Hong Kong, Vol I IMECS 2009, March 18 - 20, 2009.

[10] Takashi and Kishino, Etsuko Ueda, Yoshio Matsumoto, Masakazu Imai, Tsukasa Ogasawara," Hand Pose Estimation for Vision Based Human Interface", IEEE Transactions on Industrial Electronics, Vol.50, No.4, pp.676, 2003.

[11] Hong Li, Michael Greenspan, "Model based Segmentation and recognition of dynamic gestures in continuous video streams", Springer publications in Pattern recognition, 44(2011)1614-1628.

[12]Poizne, Zhou Ren, Junsong Yuan, Jingjing Meng, Zhengyou Zhang, "Robust Part-Based Hand Gesture Recognition Using Kinect Sensor", IEEE Transactions on Multimedia, Vol.15, No.5, August 2013.

[13]Sperling and Parish, C. Keskin, A. Erkan, L. Akarun, "Real time hand tracking and 3D gesture recognition for interactive interfaces using HMM", IEEE Transaction.

[14]Popa, M, "Hand Gesture recognition based on Accelerometer Sensors", IEEE Networked Computing and Advanced Information Management(NCM),2011.

[15]Thomas Sholmer, Benjamin Popinga, NielsHenze, Susanne Bol, "Gesture Recognition with a Wii controller", In TEI'08 Proceedings of the 2nd international conference on Tangible and embedded interaction Pages 11-14.

[16]Blanca Miriam Lee-Cosio, Carlos Delgado Mata, Jesus Ibanez, "ANN for Gesture Recognition using Accelerometer Data‖", Elsevier Publications, Procedia Technology 3 (2012) 109 – 120.

[17]Fabio Dominio, Mauro Donadeo, Pietro Zanuttigh, "Combining multiple depth-based descriptors for hand gesture recognition", Elsevier Publications, Pattern recognitions 2013.

[18]Guoliang Yang, Huan Li, Li Zhang, Yue Cao, "Research on a Skin Colour Detection Algorithm Based on Self-Adaptive Skin Colour Model", IEEE Communications and Intelligence Information Security(ICCIIS),2010.

[19]Lei Yang, Hui Li, Xiaoyu Wu, Dewei Zhao, Jun Zhai, "An algorithm of skin detection based on texture‖", IEEE Image and Signal Processing(CSIP),2011.

[20]Fabio Dominio, Mauro Donadeo, Pietro Zanuttigh, "Combining multiple depth-based descriptors for hand gesture recognition", Elsevier, Pattern recognition Letters 2013.

[21]Sheng-Yu Peng, Kanoksak Wattanachote, Hwei-Jen Lin, Kuan-Ching Li, "A Real-Time Hand Gesture Recognition System for Daily Information Retrieval from Internet", IEEE Fourth International Conference on Ubi-Media Computing, 978-0-7695-4493-9/11 © 2011.

[22]M. M. Gharasuie, H. Seyedarabi, "Real-time Dynamic Hand Gesture Recognition using Hidden Markov Models‖", 8th Iranian Conference on Machine Vision and Image Processing (MVIP), 2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP), 978-1-4673-6184-2/13/$31.00 ©2013 IEEE.

[23]Nick C. Tang, Chiou-Ting Hsu, Chih-Wen Su, Timothy K. Shih, Hong-Yuan Mark Liao, "Video in painting on Digitized Vintage Films via Maintaining Spatiotemporal Continuity", Published at IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 13, NO. 4, AUGUST 2011.

[24]Argyros, A. A., Lourakis, M. I.," Vision-Based Interpretation of Hand Gestures for Remote Control of a Computer Mouse", Proceedings of the HCI'06 workshop (in conjunction with ECCV'06), pp. 4051, 2006.

[25]Baudel, T., Beaudouin-Lafon, M. Charade, "Remote Control of Objects using Free-Hand Gestures" Communications of the ACM (36), pp. 28-35, 1993.

[26]Cassell, J., "A Framework for Gesture Generation and Interpretation", Computer Vision in Human-Machine Interaction, pp. 191-215, 1998.

[27]Elmezain, M., Al-Hamadi, A., Appenrodt, J., Michaelis B., "A Hidden Markov Model-Based Continuous Gesture Recognition System for Hand Motion Trajectory", Proceedings of Pattern Recognition, ICPR, pp. 1-4, 2008.

[28]Evoluce. "Plug Kinect in a PC a now interact with Windows 7 applications through gestures", http://www.evoluce.com/en/company/pressreleases.php?we_objectID=28, Retrieved February 15, from Evoluce, 2011.

[29]Freeman, W. T., & Roth, M., "Orientation Histograms for Hand Gesture Recognition", Proceedings of International Workshop on Automatic Face and Gesture Recognition, pp. 296-301, 1994.

[30]Freeman, W. T., Weissman, C. D., "Television Control by Hand Gestures", Proceedings of IEEE International Workshop on Automatic Face Recognition, pp. 179-183, 1995.

[31] K. Imagawa, Shan Lu, S. Igi, "Color-based hands tracking system for sign language recognition", Third IEEE International Conference on Automatic Face and Gesture Recognition, 1998.

Appendix

Screenshots

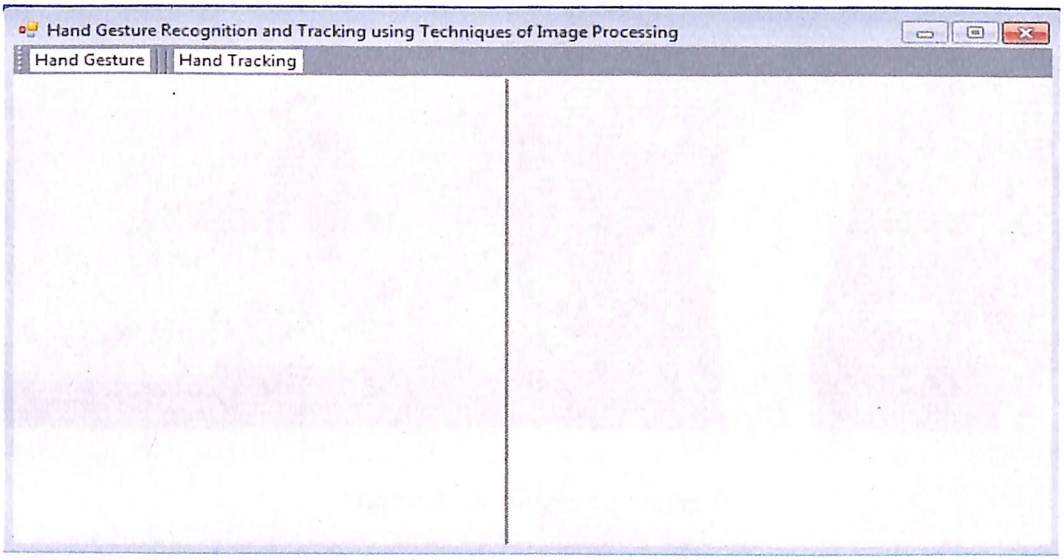


Figure A-1 The GUI of the proposed system

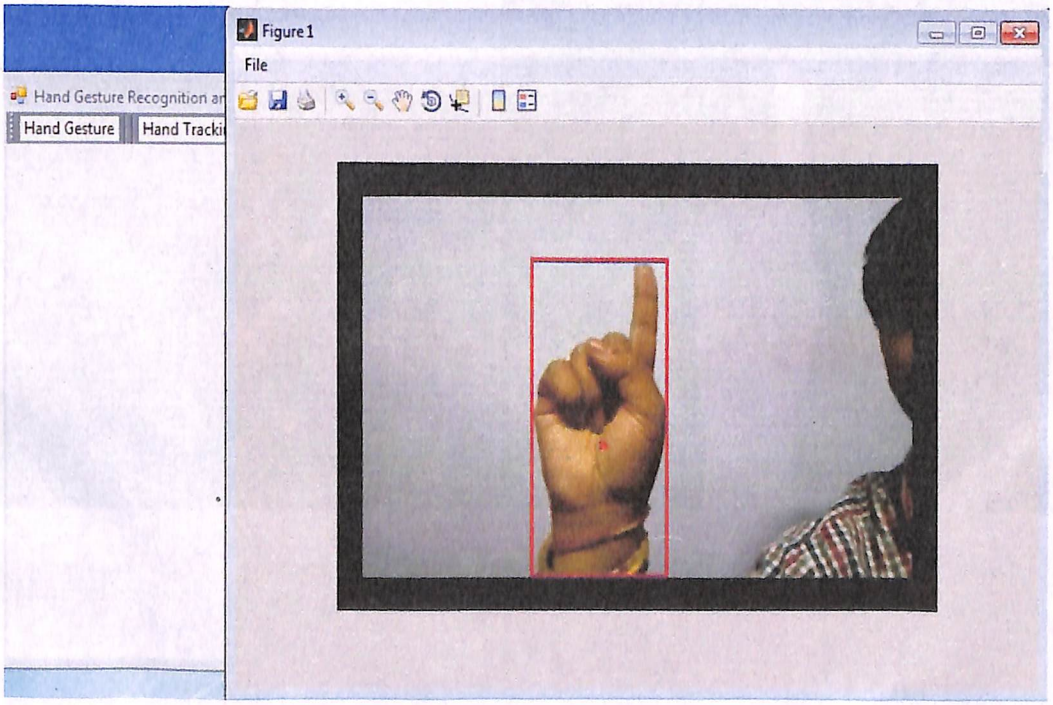


Figure A-2: Tracking the hand movements

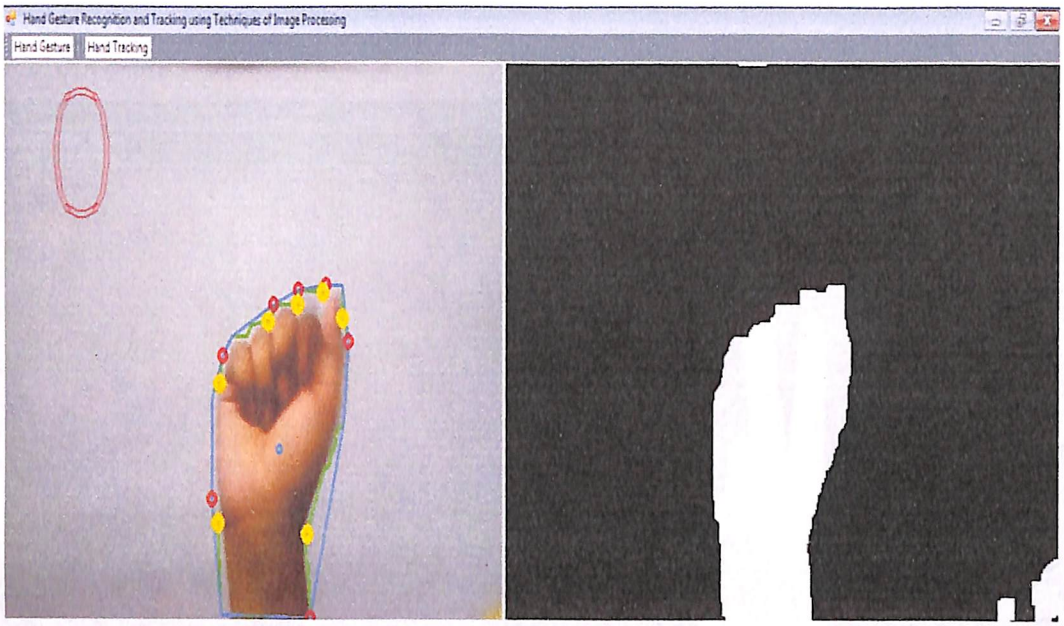


Figure A -3: Finger Counting 0

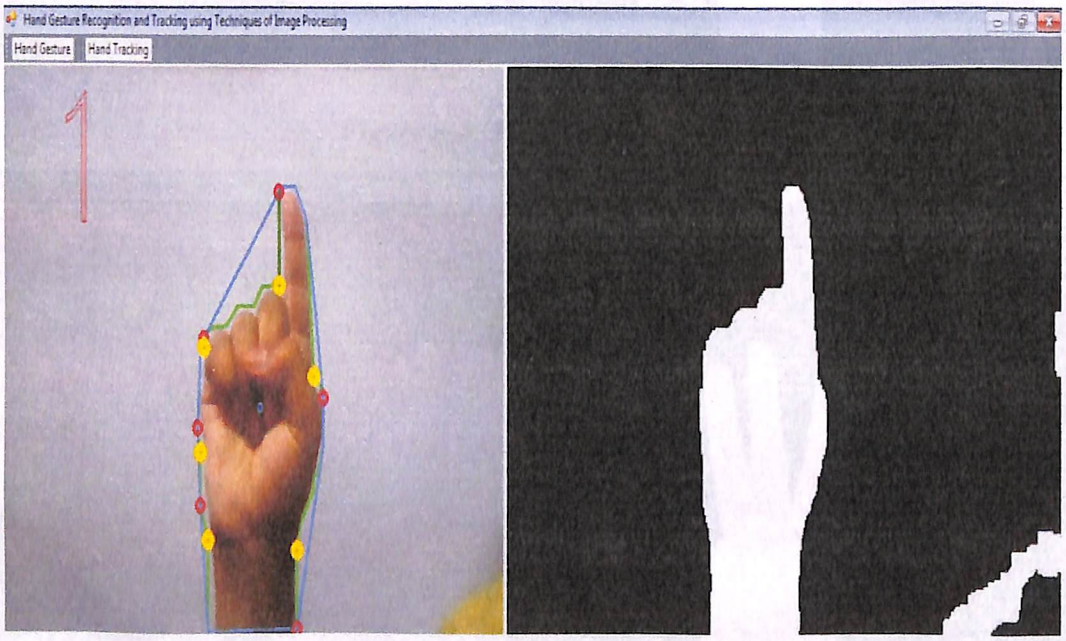


Figure A-4: Finger Counting 1

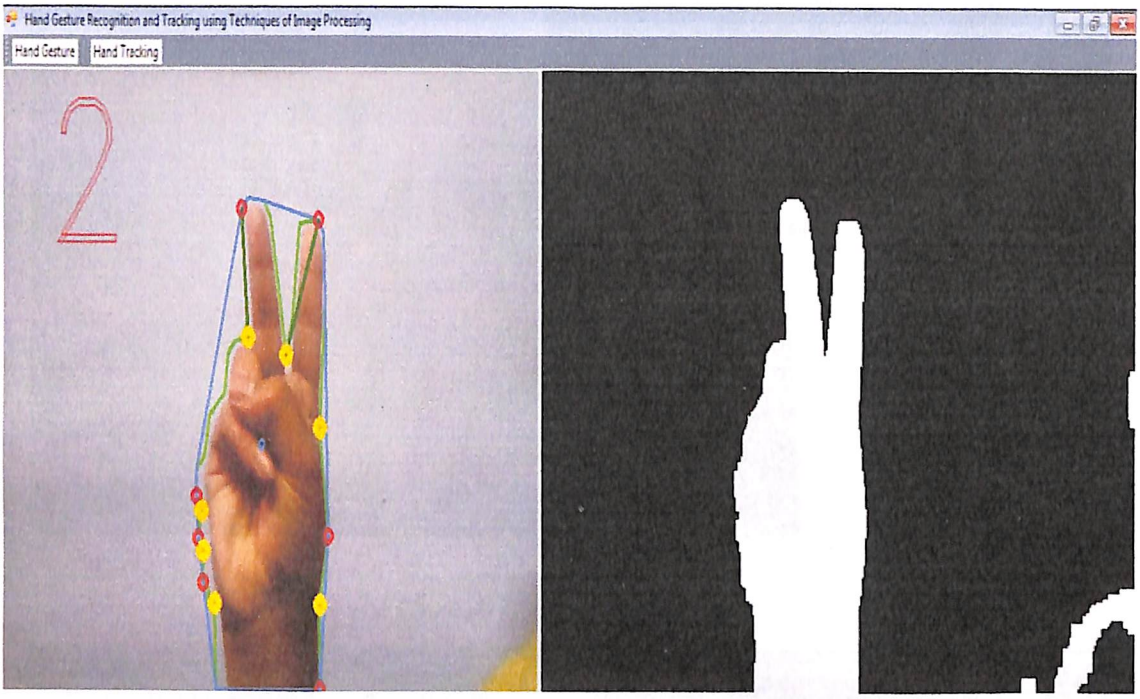


Figure A-5: Finger Counting 2

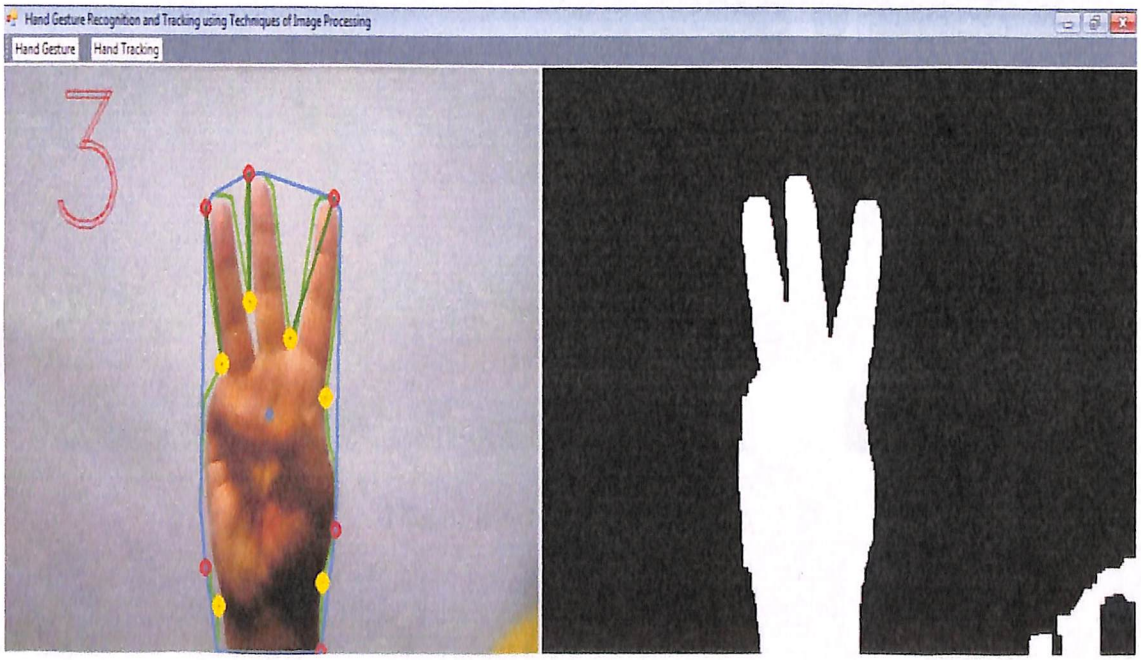


Figure A-6: Finger Counting 3

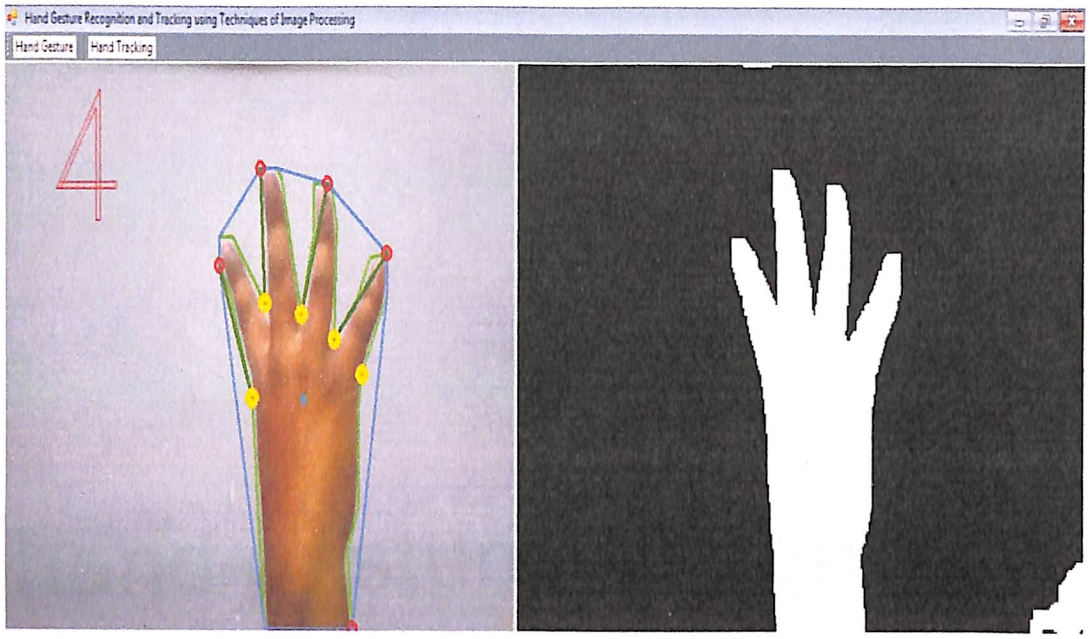


Figure A-7: Finger Counting 4

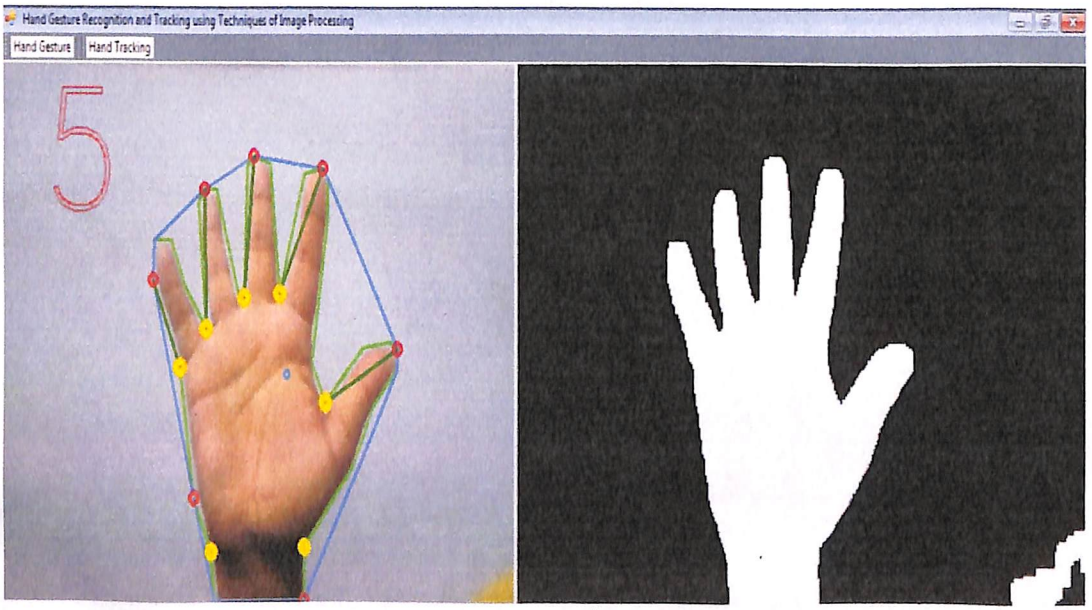


Figure A-8 Finger Counting 5

hand gesture recognistion

by Shardul Chauhan

FILE	[illegible]		
TIME SUBMITTED	02 APR [illegible]	WORD COUNT	[illegible]
SUBMISSION ID	[illegible]	CHARACTER COUNT	[illegible]

hand gesture recognistion

ORIGINALITY REPORT

20%	5%	6%	11%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Jörg Zieren. "Non-Intrusive Acquisition of Human Action", Signals and Communication Technology, 2006 Publication	5%
2	heinojorgensen.dk Internet Source	3%
3	Submitted to 14403 Student Paper	2%
4	Submitted to Higher Education Commission Pakistan Student Paper	2%
5	Submitted to National Institute of Technology, Rourkela Student Paper	2%
6	Submitted to National Institute of Technology Karnataka Surathkal Student Paper	1%
7	www.coe.utah.edu Internet Source	1%

8	Submitted to Centre for Development of Advanced Computing Student Paper	1%
9	Submitted to Monash University Student Paper	1%
10	Submitted to iGroup Student Paper	1%
11	Submitted to Universiti Tenaga Nasional Student Paper	<1%
12	kvaibhav01.wordpress.com Internet Source	<1%
13	www.pic32.ru Internet Source	<1%
14	Submitted to Texas A & M University, Kingville Student Paper	<1%
15	www.ijstr.org Internet Source	<1%
16	Submitted to Informatics Education Limited Student Paper	<1%
17	Submitted to UC, Boulder Student Paper	<1%
18	K. Srinivasan. "A new approach for human activity analysis through identification of body parts using skin colour segmentation",	<1%

	International Journal of Signal and Imaging Systems Engineering, 2010 Publication	
19	www.vision.ime.usp.br Internet Source	<1%
20	Submitted to Help University College Student Paper	<1%
21	Ding, Wenjing. "Analysis Method for Closed-Loop System", Self-Excited Vibration, 2010. Publication	<1%
22	ethesis.nitrkl.ac.in Internet Source	<1%

EXCLUDE QUOTES	OFF	EXCLUDE MATCHES	OFF
EXCLUDE BIBLIOGRAPHY	OFF		