GESTURE CONTROLLED VIRTUAL MOUSE WITH VOICE RECOGNITION

A MINI PROJECT REPORT

Submitted by

ROHITH SS	**190801065**
SAILESH BAABU S	**190801068**
SRIRAM B	**190801082**

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

IN

INFORMATION TECHNOLOGY

SRI VENKATESWARA COLLEGE OF ENGINEERING
(An Autonomous Institution; Affiliated to Anna University, Chennai -600 025)

ANNA UNIVERSITY:: CHENNAI 600 025

NOVEMBER 2022

SRI VENKATESWARA COLLEGE OF ENGINEERING

(An Autonomous Institution; Affiliated to Anna University, Chennai -600 025)

ANNA UNIVERSITY, CHENNAI – 600 025

BONAFIDE CERTIFICATE

Certified that this project report **"GESTURED CONTROLLED VIRTUAL MOUSE WITH VOICE RECOGNITION"** is the bonafide work of **"ROHITH SS(190801065), SAILESH BAABU S(190801068) , SRIRAM B (190801082) "** who carried out the project work under my supervision.

SIGNATURE	**SIGNATURE**
Dr. V. VIDHYA, M.E., Ph.D.,	**K.KIRUTHIKA DEVI, M.E.,**
HEAD OF THE DEPARTMENT	**SUPERVISOR**
	ASSOCIATE PROFESSOR
Department of Information Technology	Department of Information Technology

Submitted for the project viva-voice examination held on

INTERNAL EXAMINER **EXTERNAL EXAMINER**

SRI VENKATESWARA COLLEGE OF ENGINEERING

(An Autonomous Institution; Affiliated to Anna University, Chennai -600 025)

ANNA UNIVERSITY, CHENNAI – 600 025

BONAFIDE CERTIFICATE

Certified that this project report **"GESTURED CONTROLLED VIRTUAL MOUSE WITH VOICE RECOGNITION"** is the bonafide work of **"ROHITH SS (190801065) , SAILESH BAABU S (190801068) , SRIRAM B (190801082) "** who carried out the project work under my supervision.

SIGNATURE

NIL

INTERNAL SUPERVISOR

NIL
ASSOCIATE PROFESSOR

Department of Information Technology

SIGNATURE

NIL

EXTERNAL SUPERVISOR

Department of Information Technology

SIGNATURE
Dr. V. VIDHYA, M.E., Ph.D.,
HEAD OF THE DEPARTMENT
Department of Information Technology

Submitted for the project viva-voce examination held on_____

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ABSTRACT

In todays technological era, many technology are evolving day by day. One such promising concept is human- system interface. For example, in a stressed out mouse there is no provision to increase restriction. In wi-fi mouse, one must have Bluetooth hardware installed within the laptop and Bluetooth dongle attached. The proposed era will haven't any such boundaries and could alternatively depend on gesture recognition. In this mission, operations like clicking and dragging of objects may be achieved with different hand gestures. The proposed gadget will only require a webcam as an input tool. The software program's that will be required to put into effect the proposed machine are opencv and python. The output of the digital camera may be displayed at the machine's display screen so that it may be in addition calibrated by means of the user. The python dependencies that will be used for implementing this machine are numpy, math, wx and mouse. In this paper, we present a singular approach for human computer interplay (hci) in which cursor motion is controlled using a real-time camera.

A manner to control the Position of the cursor with the bare palms without using Any digital tool. While the operations like clicking And dragging of objects could be accomplished with special Hand gestures. The proposed gadget will handiest require a Webcam as an input tool. The software's that will be Required to put in force the proposed machine are opencv And python. The output of the camera may be displayed on The system's display screen so that it could be similarly calibrated by means of the person.

ACKNOWLEDGEMENT

We thank our Principal **Dr.S.Ganesh Vaidyanathan, Ph.D.,** Sri Venkateswara College of Engineering (Autonomous) for being the source of inspiration throughout our study in this college.

We express our sincere thanks to **Dr. Vidhya V, M.E., Ph.D.,** Head of Department, Information Technology for her permission and encouragement accorded to carry this project.

We are also thankful to **Mrs. K.Kiruthika Devi, M.E.,** project coordinator for his continual support and assistance throughout the course of this project.

With profound respect, we express our deep sense of gratitude and sincere thanks to our guide, **Mr. A.Thiyagarajan, M.E.,** Associate Professor, for her continuous and valuable guidance throughout this project.

We also express our thanks to all Faculty members, Department of Information Technology, for rendering their support.

Rohith SS

Sailesh Baabu S

Sriram.B

Table of Content

LIST OF ABBREVIATIONS

GUI	Graphical User Interface
LED	Light Emitting Diodes
HCI	Human Computer Interaction
MHI	Motion History Images
IDE	Integrated Development Environment
OpenCV	Open Source Computer Vision
HSV	Hue, Saturation, and Values
RGB	Red, Green, and Blue

Table of Figure

CHAPTER 1

1. PROJECT BACKGROUND

1.1. Introduction

A mouse, in computing terms is a pointing device that detects two-dimensional movements relative to a surface. This movement is converted into the movement of a pointer on a display that allows to control the Graphical User Interface (GUI) on a computer platform. There are a lot of different types of mouse that have already existed in the modern days technology, there's the mechanical mouse that determines the movements by a hard rubber ball that rolls around as the mouse is moved. Years later, the optical mouse was introduced that replace the hard rubber ball to a LED sensor to detects table top movement and then sends off the information to the computer for processing. On the year 2004, the laser mouse was then introduced to improve the accuracy movement with the slightest hand movement, it overcome the limitations of the optical mouse which is the difficulties to track high-gloss surfaces. However, no matter how accurate can it be, there are still limitations exist within the mouse itself in both physical and technical terms. For example, a computer mouse is a consumable hardware device as it requires replacement in the long run, either the mouse buttons were degraded that causes inappropriate clicks, or the whole mouse was no longer detected by the computer itself. Despite the limitations, the computer technology continues to grow, so does the importance of the human computer interactions. Ever

since the introduction of a mobile deve that can be interact with touch screen technology, the world is starting to demand the same technology to be applied on every technological devices, this includes the desktop system. However, even though the touch screen technology for the desktop system is already exist, the price can be very steep. Therefore, a virtual human computer interaction device that replaces the physical mouse or keyboard by using a webcam or any other image capturing devices can be an alternative way for the touch screen. This device which is the webcam will be constantly utilized by a software that monitors the gestures given by the user in order to process it and translate to motion of a points, as similar to a physical mouse.

Voice Commands:

Upcoming trending technologies such as virtual reality, augmented reality, voice interaction, IOT etc are changing the way people engage with the world and transforming digital experiences. Voice control is one of important development of human-machine interaction, which was possible because of advancement in Artificial Intelligence. In current era, we are able to train our machine to do their tasks by themselves or to think like humans using technologies like Artificial Intelligence, Machine Learning, Neural Networks, etc. we can talk to our machines with the help of virtual assistants. In recent time great appearance of voice assistants such as Apple's Siri, Google's Assistant, Microsoft's Cortana and Amazon's Alexa have been noticed due to heavy use of smartphones. Voice assistants uses technologies like voice recognition, speech synthesis, and Natural Language Processing (NLP) to provide various services which help users to perform their task using their machine by just giving commands in voice format

and also with the help of Voice Assistant there will be no need to write the commands again and again for performing particular task. Virtual assistants are very useful for old generation people, people with disabilities or special cases, small children who don't know to operate machines or smart gadgets, by making them sure that their interaction with machine is not difficult anymore and also enable them to perform Multitasking.

1.2 Problem Statement :

It's no surprised that every technological devices have its own limitations, especially when it comes to computer devices. After the review of various type of the physical mouse, the problems are identified and generalized. The following describes the general problem that the current physical mouse suffers:

- Physical mouse is subjected to mechanical wear and tear.
- Physical mouse requires special hardware and surface to operate.
- Physical mouse is not easily adaptable to different environments and its performance varies depending on the environment.
- Mouse has limited functions even in present operational environments.
- All wired mouse and wireless mouse have its own lifespan.

1.3. Motivation of Virtual Mouse

It is fair to say that the Virtual Mouse will soon to be substituting the traditional physical mouse in the near future, as people are aiming towards the lifestyle where that every technological devices can be controlled and interacted remotely without using any peripheral devices such as the remote,

keyboards, etc. it doesn't just provides convenience, but it's cost effective as well.

1.4. Convenient

It is known in order to interact with the computer system, users are required to use an actual physical mouse, which also requires a certain area of surface to operate, not to mention that it suffers from cable length limitations. Virtual Mouse requires none of it, as it only a webcam to allow image capturing of user's hand position in order to determine the position of the pointers that the user want it to be. For example, the user will be able to remotely control and interact the computer system by just facing the webcam or any other image capturing devices and moving your fingers, thus eliminating the need to manually move the physical mouse, while able to interact with the computer system from few feet away.

1.5. Cost Effective

A quality physical mouse is normally cost from the range of 30 ringgit to a hefty 400 ringgit, depending on their functionality and features. Since the Virtual Mouse requires only a webcam, a physical mouse are no longer required, thus eliminating the need to purchase one, as a single webcam is sufficient enough to allow users to interact with the computer system through it, while some other portable computer system such as the laptop, are already supplied with a built-in webcam, could simply utilize the Virtual Mouse software without having any concerns about purchasing any external peripheral devices.

1.6. Project Scope

Virtual Mouse that will soon to be introduced to replace the physical computer mouse to promote convenience while still able to accurately interact and control the computer system. To do that, the software requires to be fast enough to capture and process every image, in order to successfully track the user's gesture.

Therefore, this project will develop a software application with the aid of the latest

software coding technique and the open-source computer vision library also known as the OpenCV. The scope of the project is as below:

· Real time application.

· User friendly application.

· Removes the requirement of having a physical mouse.

The process of the application can be started when the user's gesture was captured in real time by the webcam, which the captured image will be processed for segmentation to identify which pixels values equals to the values of the defined colour. After the segmentation is completed, the overall image will be converted to Binary Image where the identified pixels will show as white, while the rest are black. The position of the white segment in the image will be recorded and set as the position of the mouse pointer, thus resulting in simulating the mouse pointer without using a physical computer mouse. The software application is compatible with the Windows platform. The functionality of the software will be coded with C++ programming language code with the integration of an external library that does the image processing known as the OpenCV.

1.7. Project Objective

The purpose of this project is to develop a Virtual Mouse application that targets a few aspects of significant development. For starters, this project aims to eliminate the needs of having a physical mouse while able to interact with the computer system through webcam by using various image processing techniques. Other than that, this project aims to develop a Virtual Mouse application that can be operational on all kind of surfaces and environment.

The following describes the overall objectives of this project:

- To design to operate with the help of a webcam.
- The Virtual Mouse application will be operational with the help of a webcam, as the webcam are responsible to capture the images in real time.
- The application would not work if there are no webcam detected.
- To design a virtual input that can operate on all surface.
- The Virtual Mouse application will be operational on all surface and indoor environment, as long the users are facing the webcam while doing the motion gesture.
- To program the camera to continuously capturing the images, which the images will be analysed, by using various image processing techniques.

As stated above, the Virtual Mouse application will be continuously capturing the images in real time, where the images will be undergo a series of process, this includes HSV conversion, Binary Image conversion, salt and pepper noise filtering, and more.

- To convert hand gesture/motion into mouse input that will be set to a particular screen position.

The Virtual Mouse application will be programmed to detect the position of the defined colors where it will be set as the position of the mouse pointers. Furthermore, a combination of different colors may result in triggering different types of mouse events, such as the right/left clicks, scroll up/down, and more.

1.8. Impact, Significance and Contribution

The Virtual Mouse application is expected to replace the current methods of utilizing a physical computer mouse where the mouse inputs and positions are done manually. This application offers a more effortless way to interact with the computer system, where every task can be done by gestures. Furthermore, the Virtual Mouse application could assist the motor-impaired users where he/she could interact with the computer system by just showing the correct combination of colors to the webcam.

1.9. Basic Fundamentals and Necessity of Voice Assistant

- Search on web
- Play a music or video
- Setting a reminder and alarm
- Run any program or application
- Getting weather updates
- Sending WhatsApp, email messages etc.

These are very few examples of tasks performed by voice assistants, we can do many more things according to our requirement. The capabilities and improvements of voice assistants are continuously developing day by day to provide better performance to users. We have used python modules and libraries for making our Desktop based voice assistant so that our personal voice assistant can run easily, smoothly on desktop.The basic idea of our Project is that the user makes a request to voice assistant through the Microphone of the device to get their work done and then their command gets converted into text. Then the text request goes to processing gives text response along with work done by voice assistant. Along with basic day to day functionalities we are also trying to implement the concept of Face detection for security purpose in our voice assistant to make it more flexible and to it make it more personal. our program uses the least amount of system resources which minimizes the expensive system requirements also reduces threat to your system as it directly does not interact with servers.

There are lots of reason why this verbal voice command application is in need in real time situations. Some of them are given below.

- **To enable a highly engaging user experience**:

Voice assistance engages users like no other interface. Users can speak to the applications naturally to ask for whatever they'd like.

- **To make application frustration free:**

We have to touch, type and mouse in the existing machine system to getting our work done, which are makes user frustrated sometimes. By using voice assistant users can directly ask what they wanted to get done.

- **To personalize your app experience for every user:**

Voice assistants are actually able to respond for every user based on their locality, language and preferences.

- **To Remove Language Barriers:**

Voice Assistant technology are blended with Translation services which helps users to handle them in their own language without concerning about language barriers which allows them to interact more freely with voice assistant.

CHAPTER 2

2.LITERATURE SURVEY

As modern technology of human computer interactions become important in our everyday lives, varieties of mouse with all kind of shapes and sizes were invented, from a casual office mouse to a hard-core gaming mouse. However, there are some limitations to these hardware as they are not as environmental friendly as it seems. For example, the physical mouse requires a flat surface to operate, not to mention that it requires a certain area to fully utilize the functions offered. Furthermore, some of these hardware are completely useless when it comes to interact with the computers remotely due to the cable lengths limitations, rendering it inaccessible.

2.1. Visual Panel

To overcome the stated problems, Zhengyou et al. (2001), proposed an interface system named Visual Panel that utilize arbitrary quadrangle-shaped planar object as a panel to allow the user to use any tip-pointer tools to interact with the computer. The interaction movements will be captured, analysed and implement the positions of the tip-pointer, resulting accurate and robust interaction with the computer. The overall system consists of panel tracker, tip-pointer tracker, holography, calculation and update, and action detector and event generator as it can simulate both mouse and keyboard.

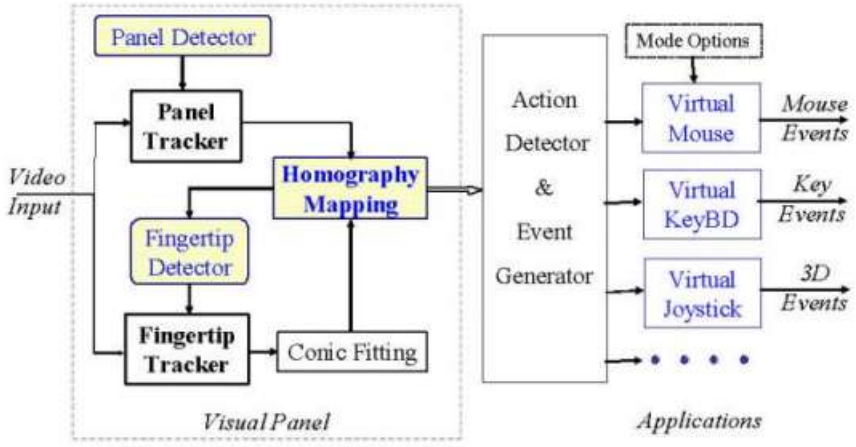


Figure 2.1: The system overview of Visual Panel (Zhengyou, Ying and Shafer, 2001)

However, although the proposed system solved the issues of cable length limitations, it still requires a certain area and material to operate. Zhengyou et al., have mentioned that the system can accepts any panel as long as it is quadrangle shaped, meaning any other shape besides stated shape are not allowed.

2.2. Virtual Mouse Using a Webcam

Another colour detection method proposed by Kazim Sekeroglu (2010), the system requires three fingers with three colour pointers to simulate the click events. The proposed system are capable of detecting the pointers by referring the defined colour information, track the motion of the pointers, move the cursor according to the position of the pointer, and simulate the single and double left or/and right click event of the mouse.

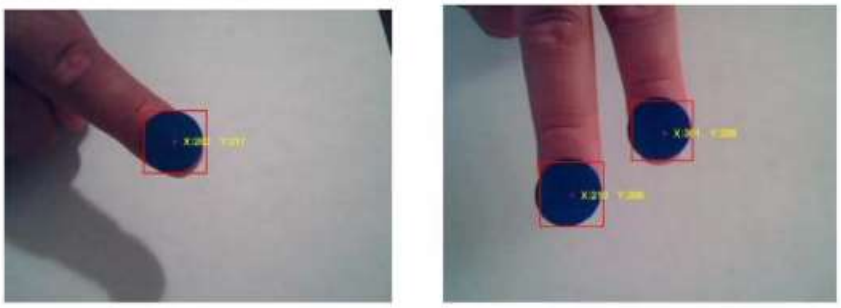


Fig 2.2 Input image using one and two fingers

To detect the colours, they have utilized the MATLAB's built in "imsubtract" function, with the combination of the noise filtering by using median filter, which are effective in filtering out or at least reduce the "salt and pepper" noise. The captured image will be converted to Binary Scale Image by using MATLAB's built in "im2bw" function to differentiate the possible values for each pixel. When the conversion is done, the captured image will undergo another filtering process by using "bwareaopen" to remove the small areas in order to get an accurate number of the object detected in the image.

2.3. Portable Vision-Based Human Computer Interaction(HCI)

Another "Ubiquitous Computing" approach proposed by Chu-Feng Lien (2015), requires only finger-tips to control the mouse cursor and click events. The proposed system doesn't requires hand-gestures nor colour tracking in order to interact with the system, instead it utilize a feature name Motion History Images(MHI) , a method that used to identify movements with a row of images in time.

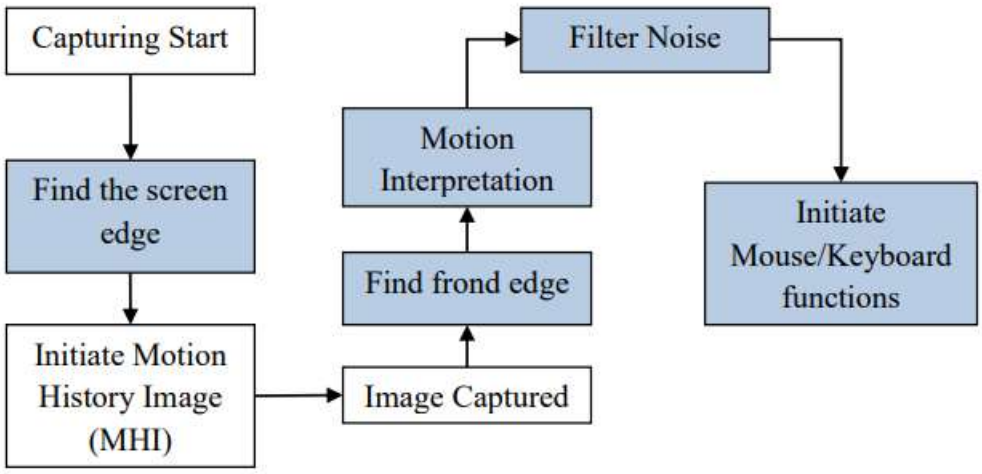


Fig 2.3 The flow chart of Portable Vision-Based Human Computer Interaction

Even though the proposed system possess good accuracy in a well-controlled environment, it does has its own limitations. The proposed system are not capable to detect fast moving movements as the frame-rates are not able to keep up, thus leading to increase of error rate. Furthermore, due to the mouse click events occurred when the finger hold on a certain positions, this may lead to user constant finger movements to prevent false alarm, which may result inconvenience.

2.4.Voice Assistant (Proton)

In today's world we train our machine to think like humans and do their task by themselves and what human being can do are being replaced by machines. Based on this situation there comes concept of voice assistant capable of completing various task for the humans based on their voice. Specific commands given by the user to virtual assistant is capable of filtering out the command and return relevant information .

People in the whole world are transforming their digital experience using

upcoming technologies like virtual reality, augmented reality, voice interaction etc. Voice control is emerging as new evolution in Human and Machine interaction where analog signal is converted by speech signal to digital wave. In Last few years huge increase in the use of smart phones led to the great use of voice assistant like Apple's Siri, Google's Assistant, Microsoft's Cortana and Amazon's Alexa etc. Voice assistants are built using technologies like voice recognition, speech synthesis, and Natural Language Processing (NLP) to provide indefinite applications to the users to make their life easy and comfortable.

Voice assistants have several interesting services for their users such as:

- Answer to questions asked by users.
- Play music from streaming music services and Playing YouTube videos.
- Set timers or alarms.
- Send WhatsApp, email messages.
- Provide information about the weather.
- Control other smart devices (lights, locks, thermostats, vacuum cleaners, switches).

The capabilities of voice assistants are continuously extending according to the users need.

According to Deepak Shende, Ria Umabiya, the AIVA (Microsoft, Google Assistant from Google, and the recently appeared intelligent assistant under the name "AIVA" 2018) aimed at developing a voice-controlled personal assistant which is doing many things such as to search the Internet. It has some new features like posting comments on the social media websites such as Facebook, Twitter, etc. By just few simple commands. You can also know the weather around you and can get the climate conditions in your region .

Tulshan explained that because of continuous typing there may be possibility

of injuries to the fingers of the user. To avoid such problems, we need to design a system in which we can get our work done through our voice commands. The voice will be recognized by the system and that recognized words will be synthesized and if they are appropriate or makes some sense then that will be printed on screen and after this again by recognizing the specific keywords the program will be compiled and executed [4]**.**

Dr. Kshama V. Kulhalli presented that survey between the top most voice assistants like Google assistant, Apple's Siri and Microsoft's Cortana. Through this survey it was concluded that Google assistant answers most accurate than others. They could understand the variations in the voice very easily [5].

2.5. Conclusion

There are abundance of methods for computer interaction besides the traditional physical mouse interaction. With the evolution of modern technology and programming, so does the Human Computer Interaction (HCI) methods, as it allows unlimited ways to access the computers. This approach allows the developers to design specific/unique system that suit the needs of the users, from gesture movement tracking to coloured tracking, it's no surprise that in near future, physical mouse will no longer be needed and be replaced by video cameras that tracks gestures.

CHAPTER 3

3.PROPOSED SYSTEM :

3.1 SYSTEM DESIGN :

During the process of colour recognition, it contains 2 major phases which are the calibration phase and recognition phase. The purpose of the calibration phase is to allow the system to recognize the Hue Saturation Values of the colours chosen by the users, where it will store the values and settings into text documents, which will be used later on during the recognition phase. While on the recognition phase, the system will start to capture frames and search for colour input with based on the values that are recorded during the calibration phase. The phases of the virtual mouse is as shown in figure below.

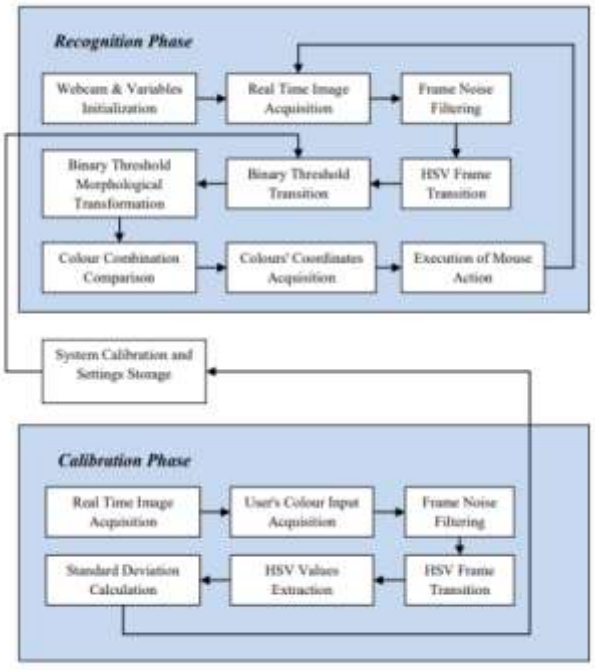


Figure 3.1: Virtual Mouse Block Diagram

3.1.1. Calibration Phase

a) Real Time Image Acquisition

The program will start of by capturing real-time images via a webcam where it will await for users' colour input. The size of the acquired image will be compressed to a reasonable size to reduce the processing loads of processing the pixels within the captured frame.

b) User's Colour Input Acquisition

The program acquires the frames that consist of input colours submitted by the users, the captured frame will be sent for process where it will undergone a series of transition and calculation to acquire the calibrated HSV values.

c) Frame Noise Filtering

Every captured frame consists of noises that will affect the performance and the accuracy of the program, therefore the frame require to be noise free. To do that, filters need to be applied on the captured frames to cancel out the unwanted noise. For the current project, Gaussian filter will be used, which is a common smoothing method to eliminate noise in a frame. This can be done by using GaussianBlur(InputArray src, OutputArray dst, Size ksize, double sigmaX, double sigmaY=0, intborderType = BORDER_DEFAULT).

Fig 3.2 Comparison between unfiltered and filtered image

d) HSV

Frame Transition

The captured frame require to be converted from a BGR format to a HSV format. Which can be done by using cvtColor(src, dst, CV_BGR2HSV).

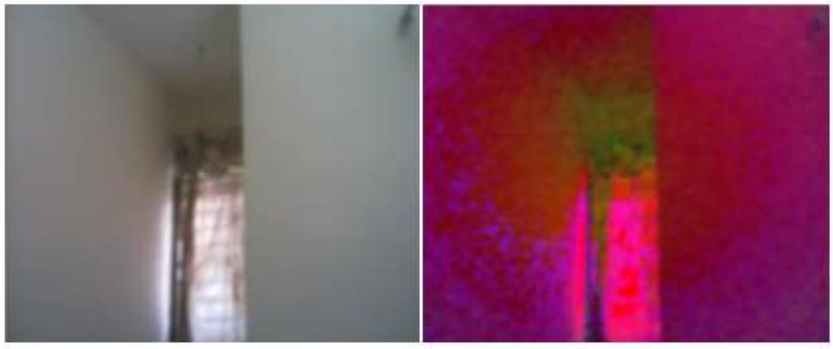


Figure 3.3: The comparison between RGB and HSV frame

e) HSV Values Extraction

In order to acquire the HSV values, the converted frame require to be split into 3 single different planes, to do that the frame needs to be divided from a multi-channel array into a single channel array, which can be done by using split(const Mat& src, Mat* mvbegin).

f) Standard Deviation Calculation

To obtain the maximum and the minimum of the HSV values, it requires to gone through the Standard Deviation calculation, a measurement used to quantify the amount of variation / dispersion among other HSV values. Furthermore, to obtain an accurate range of values, three-sigma rule are required in the calculation, so that chances of the captured values have a very high possibility to fall within the three-sigma intervals.

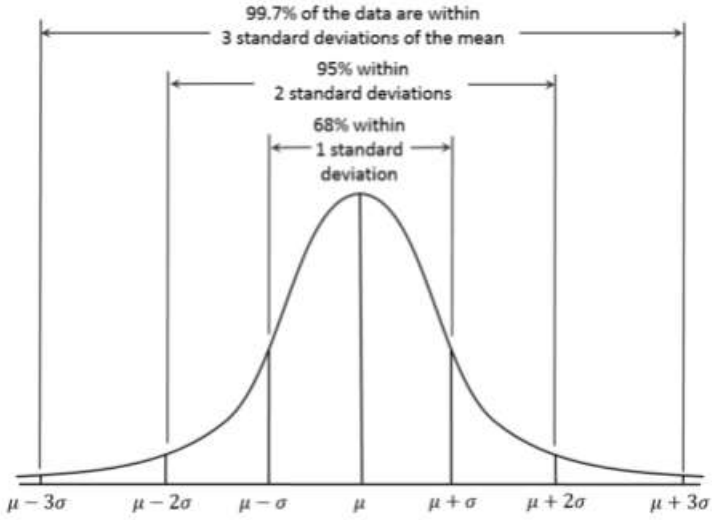


Figure 3.4: The distribution of three-sigma rule

3.1.2. Recognition Phase

a) Webcam & Variables Initialization

On the early stage of the recognition phase, the program will initialize the required variables which will be used to hold different types of frames and values where each are will be used to carry out certain task. Furthermore, this is the part where the program collects the calibrated HSV values and settings where it will be used later during the transitions of Binary Threshold.

b) Real Time Image Acquisition

The real time image is captured by using the webcam by using (cv::VideoCapture cap(0);), where every image captured are stored into a frame variable (cv::Mat), which will be flipped and compressed to a reasonable size to reduce process load.

c) Frame Noise Filtering

Similar to the noise filtering during the calibration phase, Gaussian filters will be applied to reduce the existing noise of the captured frames. This can be done by using GaussianBlur(InputArray src, OutputArray dst, Size ksize, double sigmaX, double sigmaY=0, intborderType=BORDER_DEFAULT).

d) HSV Frame Transition

The captured frame require to be converted from a BGR format to a HSV format. Which can be done by using cvtColor(src, dst, CV_BGR2HSV).

e) Binary Threshold Transition

The converted HSV frame will undergone a range check to check if the HSV values of the converted frame lies between the values of the HSV variables gathered during the calibration phase. The result of the range check will convert the frame into a Binary Threshold, where a part of the frame will set to 255 (1 bit) if the said frame lies within the specified HSV values, the frame will set to 0 (0 bit) if otherwise.

f) Binary Threshold Morphological Transformation

After the binary threshold is obtained, the frame will undergone a process called Morphological Transformation, which is a structuring operation to eliminate any holes and small object lurking around the foreground. The transformation consist of two morphological operators, known as Erosion

and Dilation. The Erosion operator are responsible for eroding the boundaries of the foreground object, decreasing the region of the binary threshold, which is useful for removing small noises. As for Dilation, it is an opposite of erosion, it increases the region of the binary threshold, allowing eroded object to return to its original form. For the current project, both operators were used for morphological Opening and Closing, where Opening consists of combination of erosion followed by dilation, which is very useful in removing noise, whereas Closing is the opposite of Opening, which is useful in closing small holes inside the foreground object.

g) Colour Combination Comparison

After obtaining results from Morphological Transformation process, the program will calculate the remaining number of objects by highlighting it as blobs, this process requires cvBlob library, which is an add-on to OpenCV. The results of the calculation will then send for comparison to determine the mouse functions based on the colour combinations found within the captured frames.

h) Colours' Coordinates Acquisition

For every object within the binary threshold, the program will highlight the overall shape of the object (cvRenderBlobs(const IplImage *imgLabel, CvBlobs &blobs, IplImage *imgSource, IplImage *imgDest, unsigned short mode=0x000f, double alpha=1.);), where it will calculate the area of the shape

and the coordinates of midpoint of the shapes. The coordinates will be saved and used later in either setting cursor positions, or to calculate the distance between two points to execute various mouse functions based on the result collected.

3.2 Hardware and Software Requirement

3.2.1. Hardware Requirement

The following describes the hardware needed in order to execute and develop

the Virtual Mouse application:

- Computer Desktop or Laptop

  The computer desktop or a laptop will be utilized to run the visual software in order to display what webcam had captured. A notebook which is a small, lightweight and inexpensive laptop computer is proposed to increase mobility.

  System will be using

  Processor : Core2Duo

  Main Memory : 4GB RAM

  Hard Disk : 320GB

  Display : 14" Monitor

- Webcam

  Webcam is utilized for image processing, the webcam will continuously

taking image in order for the program to process the image and find pixel

position.

3.2.2. Software Requirement

The following describes the software needed in-order to develop the Virtual Mouse application:

- Python Language

  The coding technique on developing the Virtual Mouse application will be the Python with the aid of the integrated development environment (IDE) that are used for developing computer programs, known as the Microsoft Visual Studio. A Python library provides more than 35 operators, covering

  basic arithmetic, bit manipulation, indirection, comparisons, logical

  operations and others.

- Open CV Library

  OpenCV are also included in the making of this program.

  OpenCV (Open Source Computer Vision) is a library of programming functions for real time computer vision. OpenCV have the utility that can read image pixels value, it also have the ability to create real time eye tracking and blink detection.

  Software will be using:

  OS : Window 7 Ultimate 64-bit

  Language : Python

Tool Used : Open CV and CMake

The software Requirements in this project include:

a. Python
b. Anaconda prompt
c. Spyder IDE
d. Modules (OpenCV, Mediapipe, Autopsy)

The MediaPipe framework is utilised for hand motion recognition and tracking, while the OpenCV library is used for computer vision.

Using MediaPipe in a machine learning pipeline requires using an open-source framework developed by Google. Cross-platform programming is made possible by the MediaPipe framework's utilisation of timeseries data. Several audio and video formats may be utilised with the MediaPipe framework. Developersuse the MediaPipe framework to create systems for application development and build and analyse systems using graphs. In a MediaPipe-based system, all steps are performed in the pipeline. Several platforms may be used for the pipeline for mobile and desktop scalability. Components of the MediaPipeframework may be broken down into the following categories: performance assessment, sensor data acquisition, and calculators. Calculators, the basic building blocks of a pipeline, are linked by streams, the medium via which data packets travel. Developers may use the graph's replacement or definition of custom calculators to construct their applications. Calculators and streams form a data flow diagram; the graph is

built using MediaPipe, and streams link the nodes.

A single-shot detector device is used to identify and recognise a hand or palm instantly. A single-shot detector is used in the MediaPipe. Since palms are simpler to train, it is initially used in the hand identification module to build a model for palm detection. Furthermore, the non-maximum suppression works better for small objects like hands or fists. In the hand area, locating knuckles or joint coordinatesis a model of a hand landmark.

Computer vision methods for object detection are part of the OpenCV library. Real-time computer vision applications may be created using OpenCV, a library for the Python programming language. As well as face and object identification, the OpenCV library is utilised in image and video processing. Because of its modular nature, OpenCV comes with several shared and four static libraries. All of these modules are available for purchase. All other modules utilise this compact module to define fundamental data structures such as Mat and the basic functions that all other modules use. Linear and non-linear image filtering, geometric transformations (resize, affine and perspective warping, general table-based remapping), colour space conversion, histograms, and so on are all included in the image processing module.

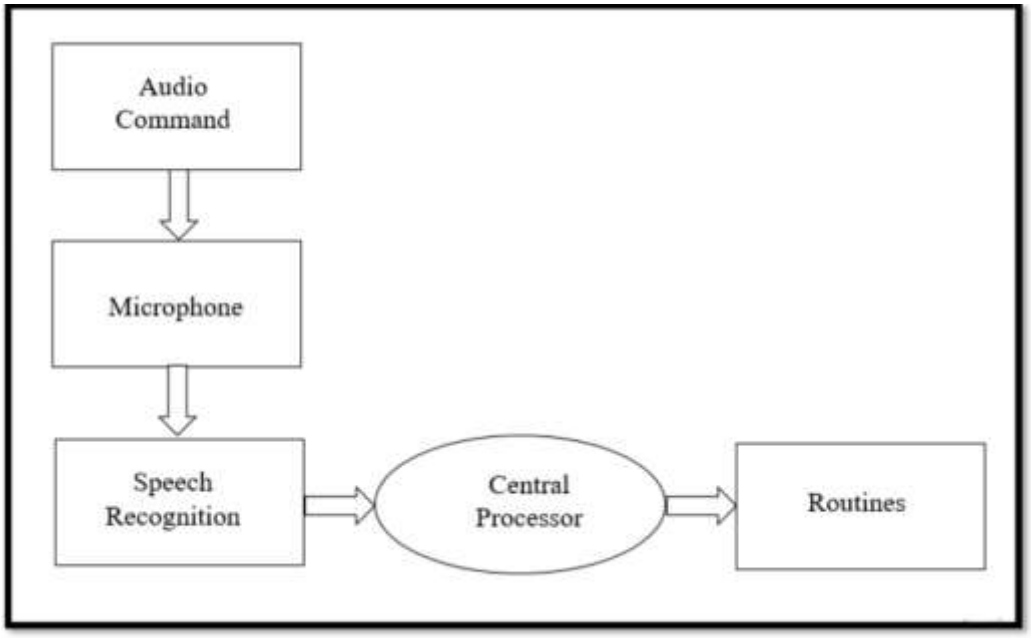
CHAPTER 4

4.METHODOLOGY :

FOR VOICE ASSISTANT

4.1 Existing System

In existing system, the audio command is taken as input through microphone of the device. The next task of voice assistant will be to analyze audio command and give appropriate output to the user. The working process of existing system is shown below



4.2

Fig 4.1 Working process

4.2. Data flow diagram (DFD):

DFD is graphical representation of system which give detail information about data flow between input and output. As level increases it elaborates detail information about data flow.

4.2.1 DFD (level 0):

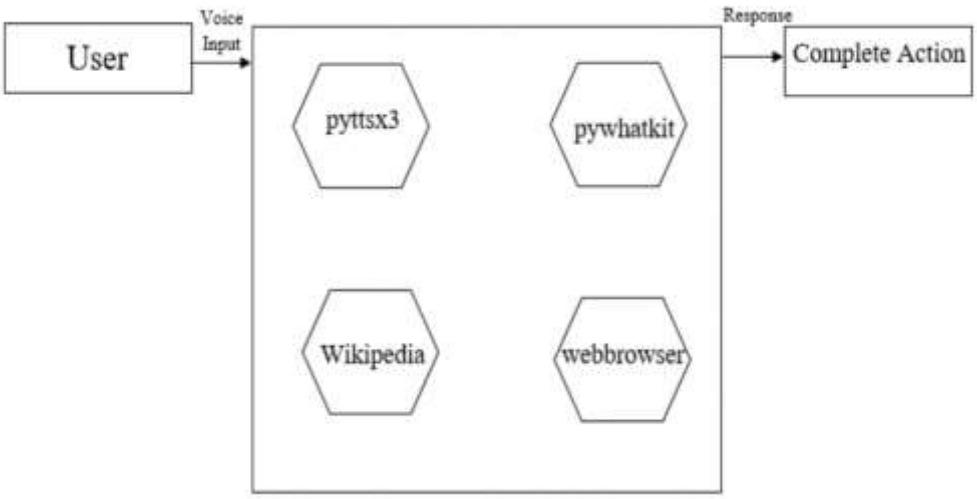


Fig 4.2 DFD(Level 0)

4.2.2 DFD (level 1):

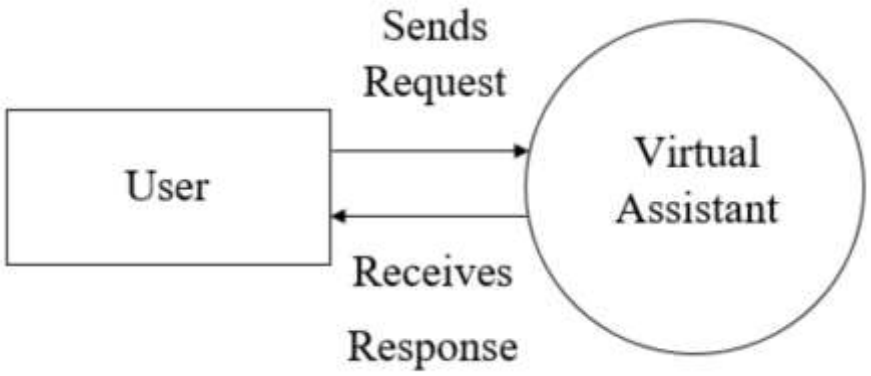


Fig 4.3 DFD**(**Level 1)

4.3. Packages used:

- **Speech Recognition:**

Speech Recognition library is used for listening to the words spoken by the users that is taken as input from microphone as a source and then process it for finding out its meaning and convert them into text format. This library allows machine system to understand the human language.

.

Pyttsx3:

Pyttsx3 stands for Python text to speech library is used for making our voice assistant talk to us. It supports common text to speech engines which is like a tool that converts text into speech and makes voice assistant able to talk to its

user. We can make it talking in both male and female voices according to requirement.

- **Wikipedia**:

We need to use Wikipedia library so that we can get information from Wikipedia on any topic or we can also ask for solution to our query or simply we can perform Wikipedia search for any topic using this library. This Library in python needs Internet connection for fetching results and it will provide results to user in text as well as voice format.

- **Datetime**:

This is an essential module to support the functionality of Date and time. Whenever user wants to know the current date and time or the user wants to schedule a task at a certain time then this module will be helpful to them.

.

PyAutoGUI:

PyAutoGUL is a Python Package which has control over the mouse and the keyboard it is able to simulate the mouse cursor moves as well as clicks the button press. With the help of particular 2-D coordinate we can click on exact location on screen.

- **PyWhatkit**:

PyWhatKit is a Python Library which has number of features like Sending messages, images through WhatsApp, playing YouTube videos, converting image to ASCII, sending emails etc.

Keyboard:

Keyboard is library in Python which provides user the full control over the Keyboard. Especially the ‘press () ’and ‘write () ’function helps for controlling keyboard keys as well as writing messages on screen.

For Gesture Control:

4.4. METHODOLOGY FOR GESTURE CONTROL

For this project we'll be using the Agile Software Development methodology approach in developing the application. The stated approach is an alternative to the traditional waterfall model that helps the project team respond to unpredictability through incremental and iterative work. It promotes adaptive planning, evolutionary development, early delivery, continuous improvement, and encourages rapid and flexible respond to change. The following describes the principles of the Agile Software Development methodology: Satisfy the customer by early and continuous delivery of workable software. Encourage changes of requirement. Workable software is delivered frequently. Continuous collaboration between the stakeholders and the developers. Project are developed around motivated individuals. Encourage informal meetings. Operational software is the principle measure of progress. Sustainable development, able to maintain a constant pace. Continuous attention to technical excellence and good design Simplicity Self-organizing teams Regular adaption to changing circumstances

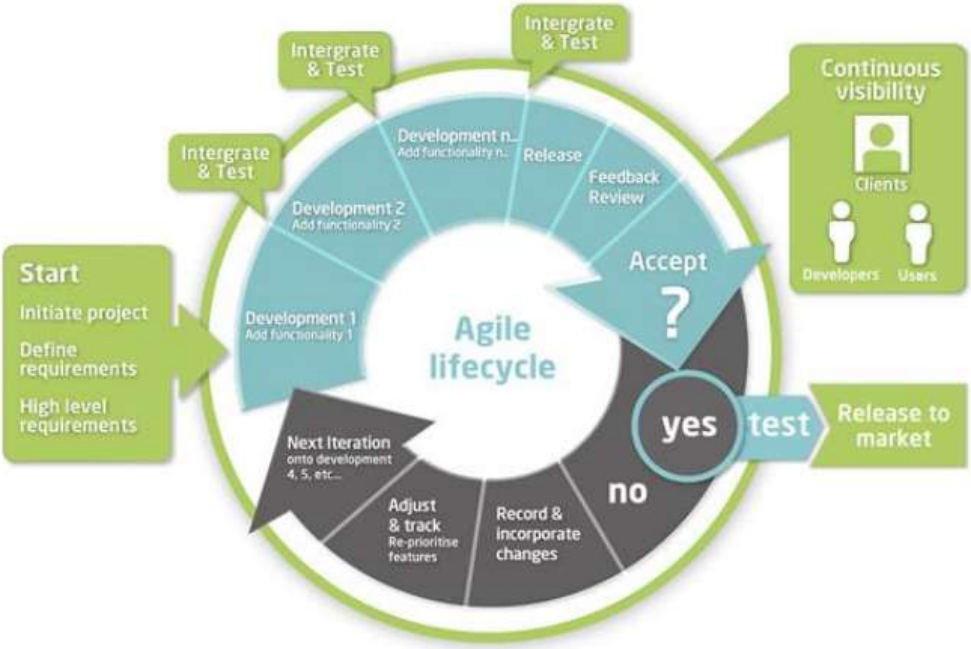


Fig 4.4 Agile methodology

The reason for choosing this methodology is due to the fact that the Virtual Mouse are still considered to be at the introduction stage, which means it still requires a great deal of extensive research and development before it could actually make it into the market. Therefore, this project requires a thorough yet iterative planning and requirements gathering where the lifecycle will be continually revisited to reevaluates the direction of the project and to eliminate the ambiguities in the process of the development, and at the same time welcome changes of requirements, which promotes adaptability and flexibility. Furthermore, due to the Virtual Mouse application are more towards serving the users, this project requires continuous customer collaboration, as they're essential for gathering the proper requirements in all aspects. This is why that the agile methodology is the ideal approach for

developing the project.

The following describes the phases within the agile methodology approach:

- Planning

A thorough planning will be conducted in this phase where the existing systems/product, for this case, physical computer mouse will be reviewed and studied to identify the problems existed, a comparison of problems will be made to compare which problems are more crucial and requires improvement. An outline objective and the scope will be identified in order to provide an alternative solution to the problem. □

- Requirement Analysis

The phase that gathers and interpreting the facts, diagnosing problems and recommending improvements to the system. In this phase, the collected problem statements will be extensively studied in order to find a proper solution or at least an improvements to the proposed system. All proposed solutions will be converted into requirements where it will be documented in a requirement specification. □

- Designing

The requirement specification from the previous phase will be studied and prioritize to determine which requirement are more important where the requirement with the highest priority will be delivered first. After the study, the system design will be prepared as it helps in defining the overall system architecture and specifying the hardware and the software requirements. □

- Building

The phase where the actual coding implementation takes place. By referring to the inputs from the system design, the system will be developed based on the prioritize requirements. However, due to we're using the agile methodology approach, the developed system will be considered as a prototype system where it will be integrated and tested by the users.

- Testing

The phase where the prototype system going through a series of test. The prototype system will first undergo integration where the features from the previous iteration cycle are added to the latest cycle. After the integration, the prototype system will be thoroughly tested by the users to determine whether they are satisfied with the latest deliverables, the completion of the project depends on whether they've accepted it or otherwise. If the users requires additional features or modification, feedback gathering will be conducted, which resulted in further modification of the requirements and features where it will recorded and documented for the requirement analysis phase on the next iteration.

CHAPTER 5

Experimental Results and Evaluation

In the proposed AI virtual mouse system, the concept of advancing the human-computer interaction using computer vision is given.

Table 1
Experimental results.

Hand tip gesture*	Mouse function performed	Success	Failure	Accuracy (%)
Tip ID 1 or both tip IDs 1 and 2 are up	Mouse movement	100	0	100
Tip IDs 0 and 1 are up and the distance between the fingers is <30	**Left button click**	99	1	99
Tip IDs 1 and 2 are up and the distance between the fingers is <40	**Right button click**	95	5	95
Tip IDs 1 and 2 are up and the distance between the fingers is >40 and both fingers are moved up the page	Scroll up function	100	0	100
Tip IDs 1 and 2 are up and the distancebetween the fingers is >40 and both fingers are moved down the page	Scroll down function	100	0	100
All five tip IDs 0, 1, 2, 3, and 4 are up	No action performed	100	0	100
Result		594	6	99

*Finger tip ID for respective fingers: tip Id 0: thumb finger; tip Id 1: index finger; tip Id 2: middle finger; tip Id 3: ring finger; tip Id 4: little finger.

Fig 5.1 Table 1

Cross comparison of the testing of the AI virtual mouse system is difficult because only limited numbers of datasets are available. The hand gestures and finger tip detection have been tested in various illumination conditions and also been tested with different distances from the webcam for tracking of the hand gesture and hand tip detection. An experimental test has been conducted to summarize the results shown in Table 1. The test was performed 25 times by 4 persons resulting in 600 gestures with manual labeling, and this test has been made in different light conditions and at different distances from the screen, and each person tested the AI virtual mouse system 10 times in normal

light conditions, 5 times in faint light conditions, 5 times in close distance from the webcam, and 5 times in long distance from the webcam, and the experimental results are tabulated in Table 1.

From Table 1, it can be seen that the proposed AI virtual mouse system had achieved an accuracy of about 99%. From this 99% accuracy of the proposed AI virtual mouse system, we come to know that the system has performed well. As seen in Table 1, the accuracy is low for “Right Click” as this is the hardest gesture for the computer to understand. The accuracy for right click is low because the gesture used for performing the particular mouse function is harder. Also, the accuracy is very good and high for all the other gestures. Compared to previous approaches for virtual mouse, our model worked very well with 99% accuracy. The graph of accuracy is shown in Figure 14.

Figure 14

Graph of accuracy.

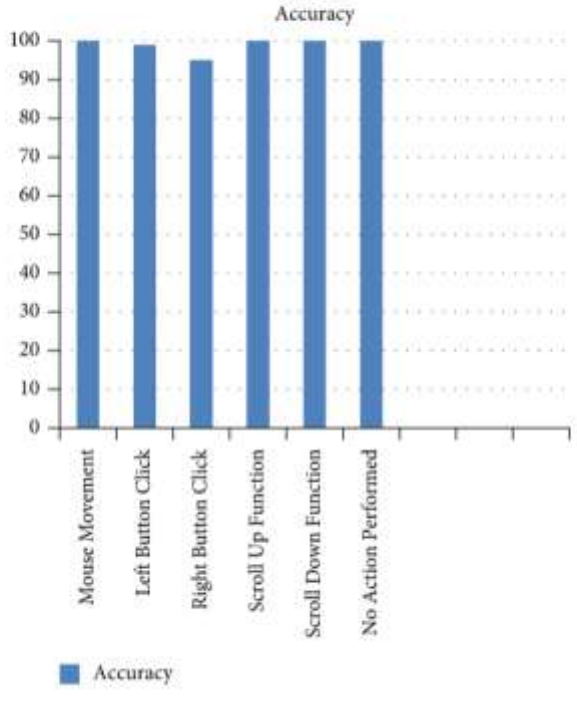


Fig 5.2 Graph 1

Table 2 shows a comparison between the existing models and the proposed AI virtual mouse model in terms of accuracy.

Table 2

Comparison with existing models.

Existing models	Accuracy (%)
Virtual mouse system using RGB-D images and fingertip detection [16]	96.13
Palm and finger recognition based [17]	78
Hand gesture-based virtual mouse [18]	78
The proposed AI virtual mouse system	99

Fig 5.3 Table 2

From Table 2, it is evident that the proposed AI virtual mouse has performed very well in terms of accuracy when compared to the other virtual mouse models. The novelty of the proposed model is that it can perform most of the mouse functions such as left click, right click, scroll up, scroll down, and mouse cursor movement using finger tip detection, and also, the model is helpful in controlling the PC like a physical mouse but in the virtual mode. Figure 15 shows a graph of comparison between the models.

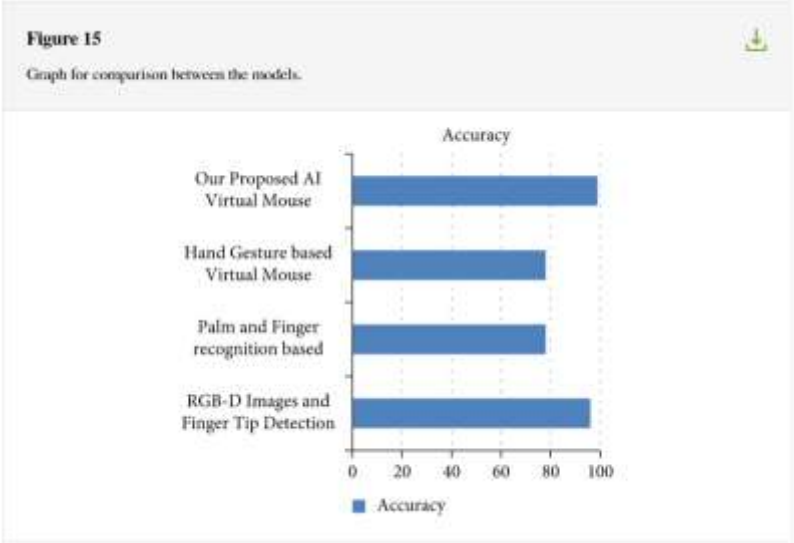


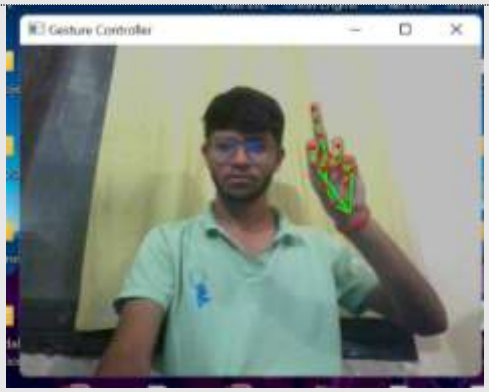
Fig 5.4 Graph 2

5.1.Performance:

The following describes the outcome of the program testing in various environments:

Gesture Controls :

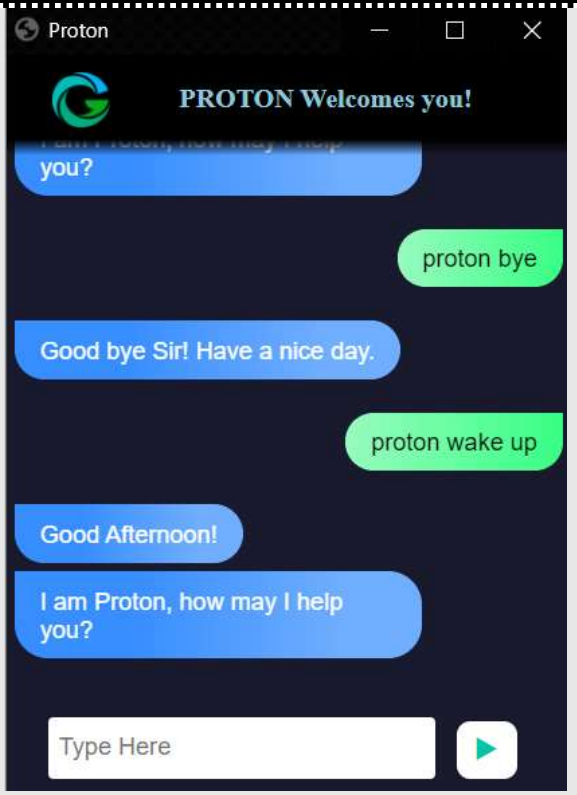
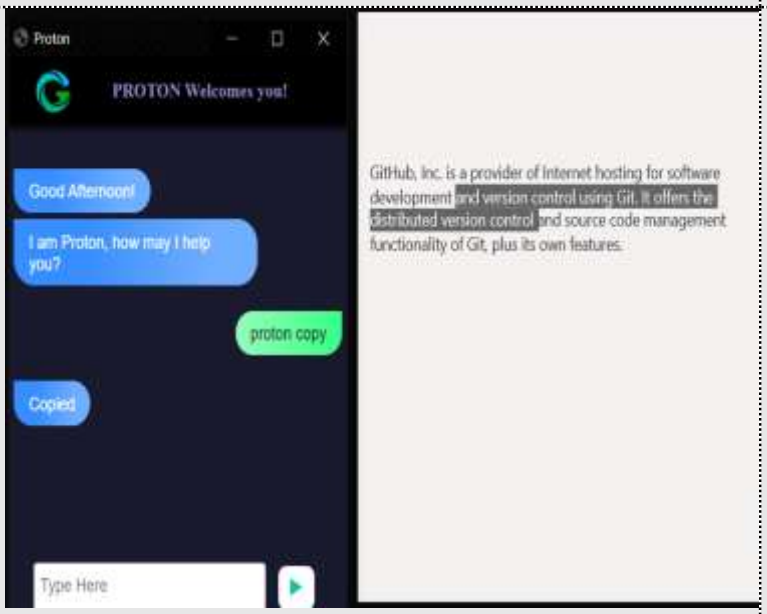
Gestures	Functions
	This shape of coordinates helps us to move cursor.

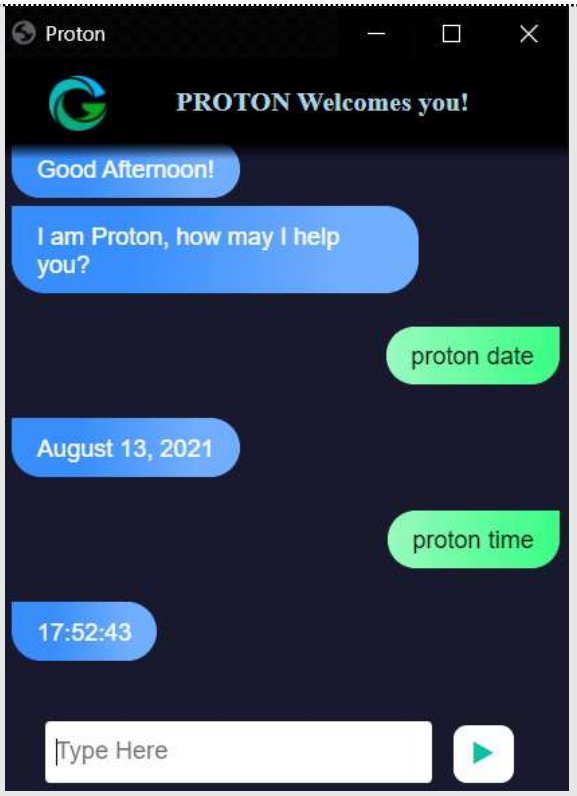
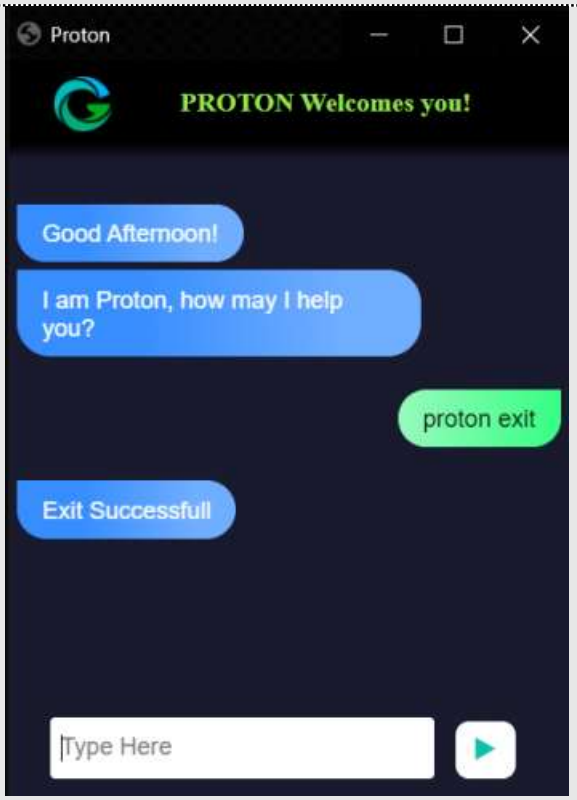
	This helps to double click the files.
	This helps us to do the function of left click.
	This function helps us to do drag files.
	This function helps us to do drop files.

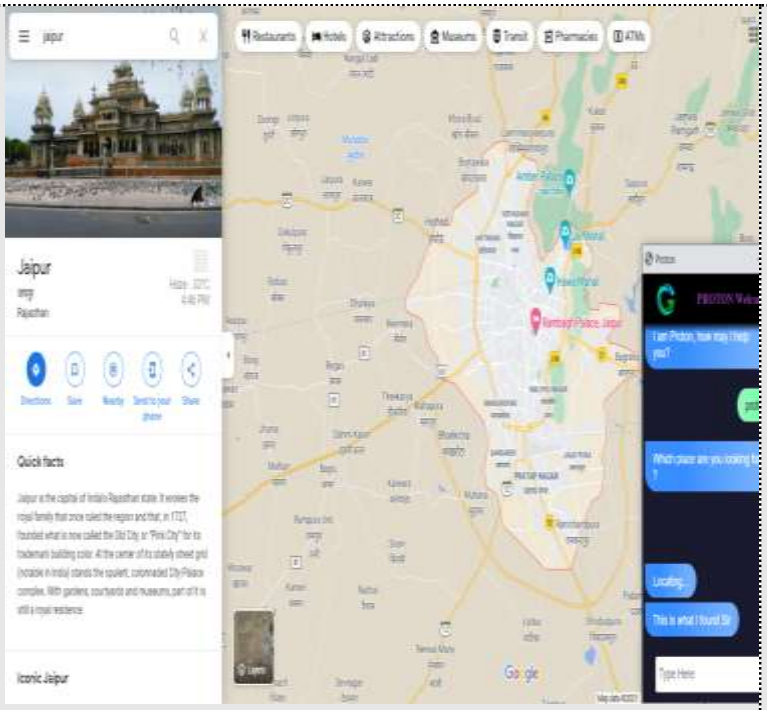
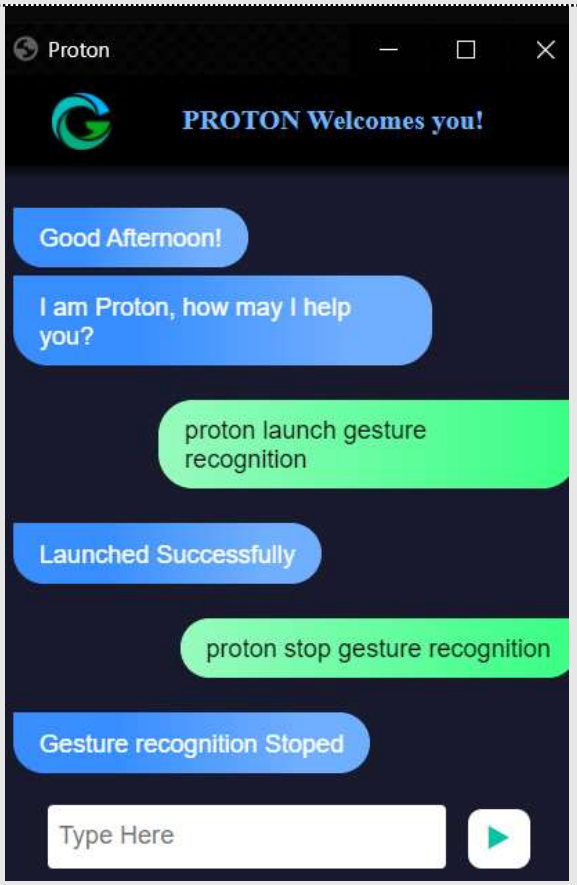
	This does the left click operation.

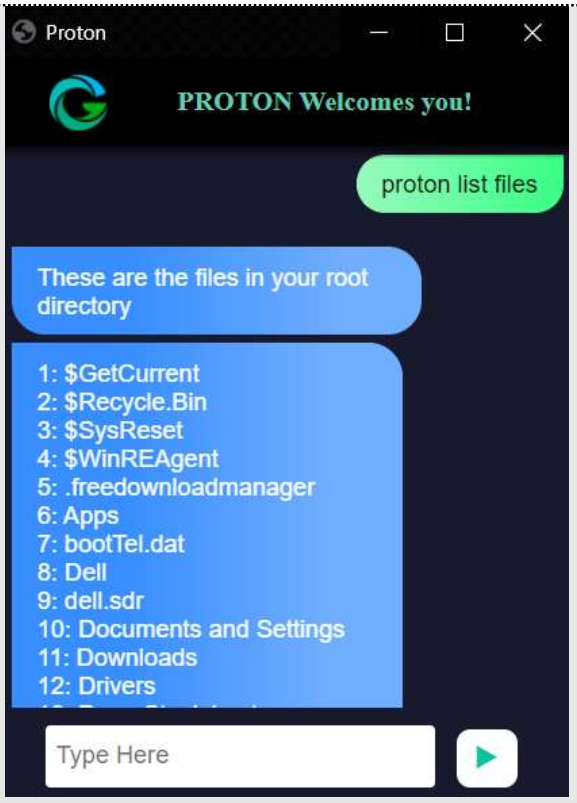
Fig 5.5 Results for gesture control

Voice Recognition:

Voice Recognition	Function
	**PROTON WAKE UP:** This command helps to turn on Proton.
	**PROTON COPY:** This command helps us to copy texts.

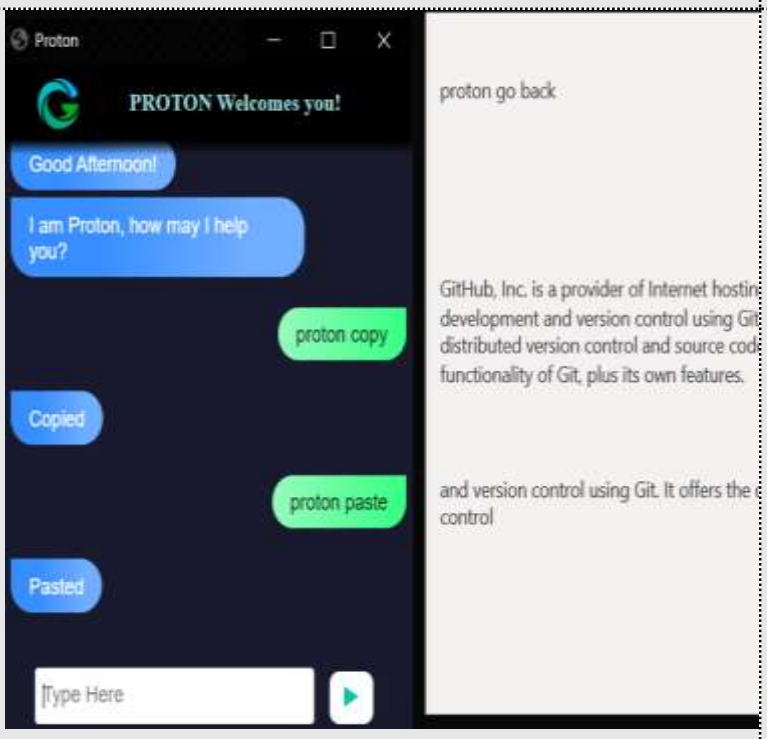
	**PROTONDATE,PROTON TIME:** This command helps us to current display time and date.
	**PROTON EXIT;** This command helps us to exit from the app.

	**PROTON FIND LOCATION:** This helps us to find the location of the given place.
	**PROTON STOP GESTURE RECOGNITION:** This command helps us to stop the gesture controls.



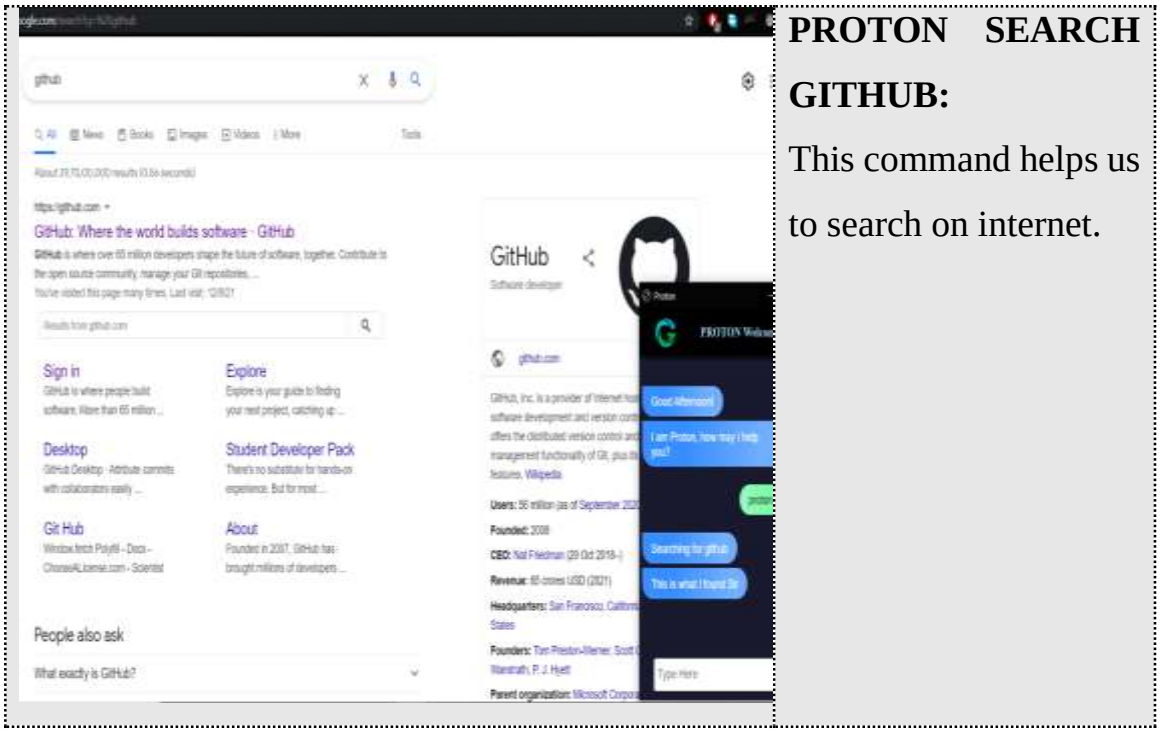
PROTON LIST FILES:

This command helps us to list the files of the root directory.



PROTON PASTE:

This command helps paste the text copied.



PROTON SEARCH GITHUB:

This command helps us to search on internet.

Fig 5.6 Results for voice control

CHAPTER 6

CONCLUTION

6.1. Overview

In conclusion, it's no surprised that the physical mouse will be replaced by a virtual non-physical mouse in the Human-Computer Interactions (HCI), where every mouse movements can be executed with a swift of your fingers everywhere and anytime without any environmental restrictions. This project had develop a colour recognition program with the purpose of replacing the generic physical mouse without sacrificing the accuracy and efficiency, it is able to recognize colour movements, combinations, and translate them into actual mouse functions. Due to accuracy and efficiency plays an important role in making the program as useful as an actual physical mouse, a few techniques had to be implemented.

First and foremost, the coordinates of the colour that are in charge of handling the cursor movements are averaged based on a collections of coordinates, the purpose of this technique is to reduce and stabilize the sensitivity of cursor movements, as slight movement might lead to unwanted cursor movements. Other than that, several colour combinations were implemented with the addition of distance calculations between two colours within the combination, as different distance triggers different mouse functions. The purpose of this implementation is to promote convenience in controlling the program without much of a hassle. Therefore, actual mouse functions can be triggered accurately with minimum trial and errors. Furthermore, to promote efficient and flexible tracking of colours, calibrations phase was implemented, this

allows the users to choose their choices of colours on different mouse functions, as long the selected colours doesn't fall within the same/similar RGB values (e.g. blue and sky-blue). Other than that, adaptive calibrations were also implemented as well, it is basically allows the program to save different set of HSV values from different angles where it will be used during the recognition phase. In Overall, the modern technologies have come a long way in making the society life better in terms of productivity and lifestyle, not the other way around. Therefore, societies must not mingle on the past technologies while reluctant on accepting changes of the newer one. Instead, it's advisable that they should embrace changes to have a more efficient, and productive lifestyle.

In this Paper we have discussed uses, methodology as well as implementation details of the personal Desktop based voice assistant using Python which is built using open-source software PyCharm as an implementation tool. This Project will be helpful for people of all generations as well as to people with some disabilities or people with some special cases. The personal voice assistant will be easy to use and will reduce the manual human efforts for performing various tasks. The functionality of the current voice assistant system is limited to working on Desktop based and working online (required to have internet connection to perform tasks) only. The voice assistant system is modular in nature so that addition of new features is possible without disturbing current system functionalities.

6.2. References:

1. Rao, A.K., Gordon, A.M., 2001. Contribution of tactile information to accuracy in pointing movements. Exp. Brain Res. 138, 438–445. https://doi.org/10.1007/s002210100717.

2. Masurovsky, A., Chojecki, P., Runde, D., Lafci, M., Przewozny, D., Gaebler, M.,2020. Controller-Free Hand Tracking for Grab-and Place Tasks in Immersive Virtual Reality: Design Elements and Their Empirical Study. Multimodal Technol.Interact.4,91. https://doi.org/10.3390/mti4040091.

3. Lira, M., Egito, J.H., Dall'Agnol, P.A., Amodio, D.M., Gonçalves, Ó.F., Boggio, P.S., 2017. The influence of skin colour on the experience of ownership in the rubber hand illusion. Sci. Rep. 7, 15745. https://doi.org/10.1038/s41598-017- 16137-3.

4. Inside Facebook Reality Labs: Wrist-based interaction for the next computing platform [WWW Document], 2021 Facebook Technol. URL https://tech.fb.com/inside-facebook-realitylabs-wrist-basedinter action-for-the-next computing-platform/ (accessed 3.18.21).

5. Danckert, J., Goodale, M.A., 2001. Superior performance for visually guided pointing in the lower visual field. Exp. Brain Res. 137, 303–308. https://doi.org/10.1007/s002210000653.

6. Carlton, B., 2021. HaptX Launches True-Contact Haptic Gloves For VR And Robotics. VRScout. URL https://vrscout.com/news/haptx-truecontact-haptic-gloves-vr/ (accessed 3.10.21).

7. Brenton, H., Gillies, M., Ballin, D., Chatting, D., 2005. D.: The uncanny valley: does it exist, in: In: 19th British HCI Group Annual Conference: Workshop on Human-Animated Character Interaction.

8. Buckingham, G., Michela kakis, E.E., Cole, J., 2016. Perceiving and acting upon weight illusions in the absence of somatosensory information. J. Neurophysiol. 115, 1946–1953. https://doi.org/10.1152/jn.00587.2015.

9. J. Katona, "A review of human–computer interaction and virtual reality research fields in cognitive Info Communications," Applied Sciences, vol. 11, no. 6, p. 2646, 2021.View at: Publisher Site | Google Scholar.

10. D. L. Quam, "Gesture recognition with a DataGlove," IEEE Conference on Aerospace and Electronics, vol. 2, pp. 755–760, 1990. View at: Publisher Site | Google Scholar.

11. D.-H. Liou, D. Lee, and C.-C. Hsieh, "A real time hand gesture recognition system using motion history image," in Proceedings of the 2010 2nd International Conference on Signal Processing Systems,

IEEE, Dalian, China, July 2010. View at: Publisher Site | Google

12. S. U. Dudhane, "Cursor control system using hand gesture recognition," IJARCCE, vol. 2, no. 5, 2013. View at: Google Scholar.

13. K. P. Vinay, "Cursor control using hand gestures," International Journal of Critical Accounting, vol. 0975–8887, 2016. View at: Google Scholar.

14. L. Thomas, "Virtual mouse using hand gesture," International Research Journal of Engineering and Technology (IRJET, vol. 5, no. 4, 2018. View at: Google Scholar.

15. P. Nandhini, J. Jaya, and J. George, "Computer vision system for food quality evaluation—a review," in Proceedings of the 2013 International Conference on Current Trends in Engineering and Technology (ICCTET), pp. 85–87, Coimbatore, India, July 2013. View at: Publisher Site | Google Scholar.

16. J. Jaya and K. Thanushkodi, "Implementation of certain system for medical image diagnosis," European Journal of Scientific Research, vol. 53, no. 4, pp. 561–567, 2011. View at: Google Scholar.

6.3. Future Works

There are several features and improvements needed in order for the program to be more user friendly, accurate, and flexible in various environments. The following describes the improvements and the features required:

a) Smart Recognition Algorithm Due to the current recognition process are limited within 25cm radius, an adaptive zoom-in/out functions are required to improve the covered distance, where it can automatically adjust the focus rate based on the distance between the users and the webcam.

b) Better Performance The response time are heavily rely on the hardware of the machine, this includes the processing speed of the processor, the size of the available RAM, and the available features of webcam. Therefore, the program may have better performance when it's running on a decent machine with a webcam that performs better in different types of lightings.