**HAND GESTURE AND**

**VOICE CONTROLLED ASSISTANT**

***A Report submitted***

***in partial fulfilment for the Degree of***

**B. Tech in**

**Computer Engineering *by***

**SAISH KADUSKAR**

**SHANTANU KADAM**

**RAJ DESHMUKH**

**OMKAR KHANDAGALE**

pursued in

**Department of Computer Engineering**

**Vishwakarma University**

To



**VISHWAKARMA UNIVERSITY**

**PUNE**

**DECEMBER, 2021**

**CERTIFICATE**

This is to certify that the project report entitled **HAND GESTURE RECOGNITION AND VOICE CONTROL** submitted by **Saish Kaduskar, Shantanu Kadam, Raj Deshmukh, Omkar Khandagale** to the Department of Computer Engineering, Science Technology, Pune, in partial fulfilment for the award of the degree of **B. Tech in Computer Engineering** is a *bona fide* record of project work carried out by him under my/our supervision. The contents of this report, in full or in parts, have not been submitted to any other Institution or University for the award of any degree or diploma.

|  |  |
| --- | --- |
| <Signature> | <Signature> |
| Dr. Prasad Gokhale | Prof. Shailesh Thaware |
| Supervisor | Co-Supervisor |
| Department of  Computer Engineering | Department of  Computer Engineering |
|  |  |

Pune Counter signature of HOD with seal

December, 2021

**DECLARATION**

I declare that this project report titled **HAND GESTURE RECOGNITION AND VOICE CONTROL** submitted in partial fulfilment of the degree of **B. Tech in Computer Engineering** is a record of original work carried out by me under the supervision of **Dr. Prasad Gokhale**, and has not formed the basis for the award of any other degree or diploma, in this or any other Institution or University. In keeping with the ethical practice in reporting scientific information, due acknowledgements have been made wherever the findings of others have been cited.

SAISH KADUSKAR (201900446)

SHANTANU KADAM (201900851)

RAJ DESHMUKH (201900450)

OMKAR KHANDAGALE (201901276)

Pune-411056

24 Dec 2021

**ACKNOWLEDGMENTS**

In successfully completing this project, many people have helped me. I would like to thank all those who are related to this project.

Primarily, I would thank God for being able to complete this project with success. Then I will thank my teacher **Dr. PRASAD GOKHALE**, under whose guidance I learned a lot about this project. Her suggestions and directions have helped in the completion of this project.

Finally, I would like to thank my friends who have helped me with their valuable suggestions and guidance and have been very helpful in various stages of project completion.

SAISH KADUSKAR (201900446)

SHANTANU KADAM (201900851)

RAJ DESHMUKH (201900450)

OMKAR KHANDAGALE (201901276)

**ABSTRACT**

The established way of interfacing with most computer systems is a mouse and keyboard. Hand gestures and Voice Commands are an intuitive and effective touchless way to interact with computer systems. Gesture recognition has a wide area of application including human machine interaction, sign language, etc. However, hand gesture-based systems have seen low adoption among end-users primarily due to numerous technical hurdles in detecting in-air gestures accurately. The purpose of the implemented software is to control the computer system using hand gesture and voice commands.

The apparition on market of the low-cost webcams and microphones with, at least, satisfactory qualities open up new directions regarding the implementation of human computer interaction (HCI) interface. This software enables users to interact in real-time with computers having only RGB cameras and microphone.

**TABLE OF CONTENTS**

**DESCRIPTION PAGE NUMBER**

CERTIFICATE 2

DECLARATION 3

ACKNOWLEDGEMENTS 4

ABSTRACT 5

LIST OF FIGURES 8

LIST OF TABLES 8

1. INTRODUCTION 9

1.1 Desktop Assistant

1.1.1 Motivation of Study

1.1.2 Objectives

1. LITERATURE SURVEY 11
2. EXISTING SYSTEM 12

3.1 Existing System

3.2 Drawbacks of Existing system

3.3 Problem Statement

1. REQUIREMENTS 14

4.1 Requirements for Development

Hardware Requirements

Software Requirements

4.2 Requirements for Users

Hardware Requirements

Software Requirements

4.3 User Requirements

5. PROJECT DESIGN 16

5.1 Class Diagram

5.2 Sequence Diagram

5.3 High level diagram

6. FUTURE SCOPE 20

6.1 Future Scope

7. CONCLUSION 21

8. REFERENCE 22

**LIST OF FIGURES**

**FIGURE TITLE PAGE NUMBER**

5.1 Class Diagram 16

5.2 High Level Diagram 24

**LIST OF TABLES**

**TABLE TITLE PAGE NUMBER**

4.1 Hardware Requirement 14

4.2 Software Requirement 14

5.1 CRC 1 17

5.2 CRC 2 17

5.3 CRC 3 18

5.4 CRC 4 18

**CHAPTER 1**

**INTRODUCTION**

**1.1. Computer Control**

Computer technology has tremendously grown over the past decade and has become a necessary part of everyday live. The primary computer accessory for Human Computer Interaction (HCI) is the mouse. The mouse is not suitable for HCI in some real-life situations, such as with Human Robot Interaction (HRI). There have been many researches on alternative methods to the computer mouse for HCI. The most natural and intuitive technique for HCI, that is a viable replacement for the computer mouse is with the use of hand gestures and voice commands. This project is therefore aimed at investigating and developing a Computer Control (CC) system using hand gestures and voice commands. Most laptops today are equipped with microphones and webcams, which have recently been used in security applications utilizing face recognition.

In order to harness the full potential of a webcam, it can be used for vision-based CC, which would effectively eliminate the need for a computer mouse or mouse pad. Over the past decades there have been significant advancements in HCI technologies for gaming purposes, such as the Microsoft Kinect and Nintendo Wii. These gaming technologies provide a more natural and interactive means of playing videogames. Motion controls is the future of gaming and it have tremendously boosted the sales of video games, such as the Nintendo Wii which sold over 50 million consoles within a year of its release. HCI using hand gestures is very intuitive and effective for one-to-one interaction with computers and it provides a Natural User Interface (NUI). There has been extensive research towards novel devices and techniques for cursor control using hand gestures. Besides HCI, hand gesture recognition is also used in sign language recognition, which makes hand gesture recognition even more significant

**1.1.1 Motivation of the study**

We had chosen this project with an interest of learning the direct interaction of humans with the consumer electronic devices. This takes the user experience to a whole new level. The gesture control technology would reduce our dependence on the age-old peripheral devices hence it would reduce the overall complexity of the system. Initially this technology was considered in the field of gaming (like Xbox Kinect), but the application of motion/gesture control technology would be more diverse if we apply it to our other electronics like computers, televisions, etc., for our day-to-day purposes like scrolling, selecting, clicking etc.

* + 1. **Objectives**
* Develop a Computer Control system to detect, recognize and interpret the hand gestures through computer vision for Windows Operating System.
* Helps out to do other task with the help of Hand and Voice.
* Develop an application which can be used by those people who have less knowledge in how to interact with system.
* Help disabled individual to interact with computer systems.

**CHAPTER 2**

**LITERATURE SURVEY**

**2.1 Review of Literature**

* Hand and Voice Recognition is blessing for everyone in this new era of 21st century.
* By using a virtual assistant, businesses only have to pay for work when they need it. They don’t have to worry about finding something to keep an employee busy if there isn’t enough work.
* It has paved way for a new technology where we can interact with machine as people do with humans.
* This new technology will attract almost everyone through interaction with system.
* It makes day to day operation on your desktop more reliable & easier.
* In some areas, the demand for gesture and voice control is extremely high.

**CHAPTER 3**

**EXISTING SYSTEM**

**3.1 Existing System**

There are various software available in the market for gesture recognition/control and Virtual assistant. Some are Virtual (Digital) assistant and some are Assistant devices. Virtual Assistant such as Cortana for windows 10, SIRI for Apple devices, Google Assistant for Android smartphones, and some Devices like ALEXA.

**SIRI by Apple**

The Most well-known utilization of iPhone is “SIRI”. It is named as Personal Assistant with Voice Recognition Intelligence. Siri supports a wide range of user commands, including performing phone actions, checking basic information, scheduling events and reminders, handling device settings, searching the Internet, etc.

**Cortana by Windows**

Cortana is Microsoft’s personal productivity assistant that helps you save time and focus attention on what matters most. Cortana can set reminders, recognize natural voice without the requirement for keyboard input, and answer questions using information from the Bing search engine (e.g., current weather and traffic conditions, sports scores)

**Eva facial mouse**

Android application for cursor controlling using head movement**.**

**Gesture Control**

Android application to perform swipe operation using hand gesture.

**3.1.1 Drawbacks of existing systems**

* SIRI only works on apple devices and does not provide any type of gesture control
* Cortana only works on Windows 10 and does not provide any type of gesture control
* Only swipe function is available in Gesture Control android application
* Eva Facial Mouse - Only cursor control operation is available.

**3.2 Problem Statement**

We are developing software to operate the computer system using hand gestures and voice commands opposed than the traditional technique of mouse and keyboard after analysing the shortcomings of several existing systems.

**CHAPTER 4**

**REQUIREMENTS**

**4.1 Requirements for Development**

**4.1.1 Hardware Requirements**

|  |  |
| --- | --- |
| **Processor** | Intel(R) Core (TM) i7-8565U CPU @ 1.80GHz 1.99 GHz |
| **RAM** | 16.0 GB |
| **Other Hardware’s** | Webcam, Headphone/Microphone/Speaker |

Table No. 4.1 Hardware Requirements

**4.1.2 Software Requirements**

|  |  |
| --- | --- |
| **Operating System** | Windows 10 |
| **Front End** | Python |
| **Back End** | Python |
| **Browser** | Google Chrome |
| **Editor** | PyCharm Community Edition |
| **Internet Connection** | Yes |

Table No. 4.2 Software Requirements

**4.2 Requirements for User**

**4.2.1 Hardware Requirements**

**Processor:** Intel i3 and above

**RAM:** 4.00 GB

**Free Storage:** 200 MB

**Other Requirements:** Webcam, Microphone, Speaker

**4.2.2 Software Requirements**

**Operating System:** Windows 7 and above

**Browser:** Google Chrome

**Active Internet Connection:** 1MBPS and above

**4.3 Other Requirements**

* A clear white background
* There should be no other objects (specially skin-coloured objects) in front of the webcam other than hand
* User must know English language.
* User must be around the system.

**CHAPTER 5**

**PROJECT DESIGN**

**5.1 Class diagram: -**

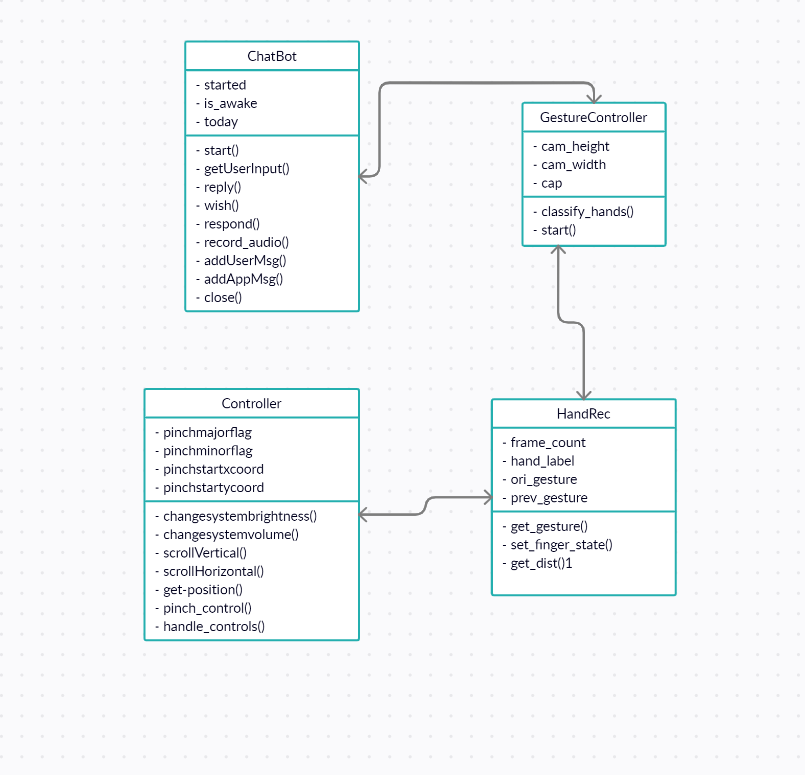


Figure 5.1 Class Diagram

CRC 1

|  |  |
| --- | --- |
| Class Name | ChatBot |
| Variables | * Started * Today * Is\_awake |
| Services | * Start() * GetUserInput() * Reply() * Close() |
| Responsibilities | 1. Start the Chatbot 2. Get user input and reply accordingly 3. Close chatbot |

Table No. 5.1

CRC 2

|  |  |
| --- | --- |
| Class Name | GestureController |
| Variables | * Cam\_height * Cam\_width |
| Services | * Classify\_hands() * Start() |
| Responsibilities | 1. Start camera and take video input 2. Detect hand and keep tracking |

Table No. 5.2

CRC 3

|  |  |
| --- | --- |
| Class Name | HandRec |
| Variables | * Frame\_count * Hand\_label |
| Services | * Get\_gesture() * Set\_finger\_state() * Get\_dist() |
| Responsibilities | Detect gestures |

Table No. 5.3

CRC 4

|  |  |
| --- | --- |
| Class Name | Controller |
| Variables | * Pinchmajorflag * Pinchminorflag * Pinchstartxcoord * pinchstartycoord |
| Services | * -changesystembrightness() * changesystemvolume() * scrollVertical()scrollHorizontal() * Get\_position() * pinch\_control() * handle\_controls() |
| Responsibilities | Perform operations according detected gestures |

Table No. 5.4

**5.3 High level diagram**

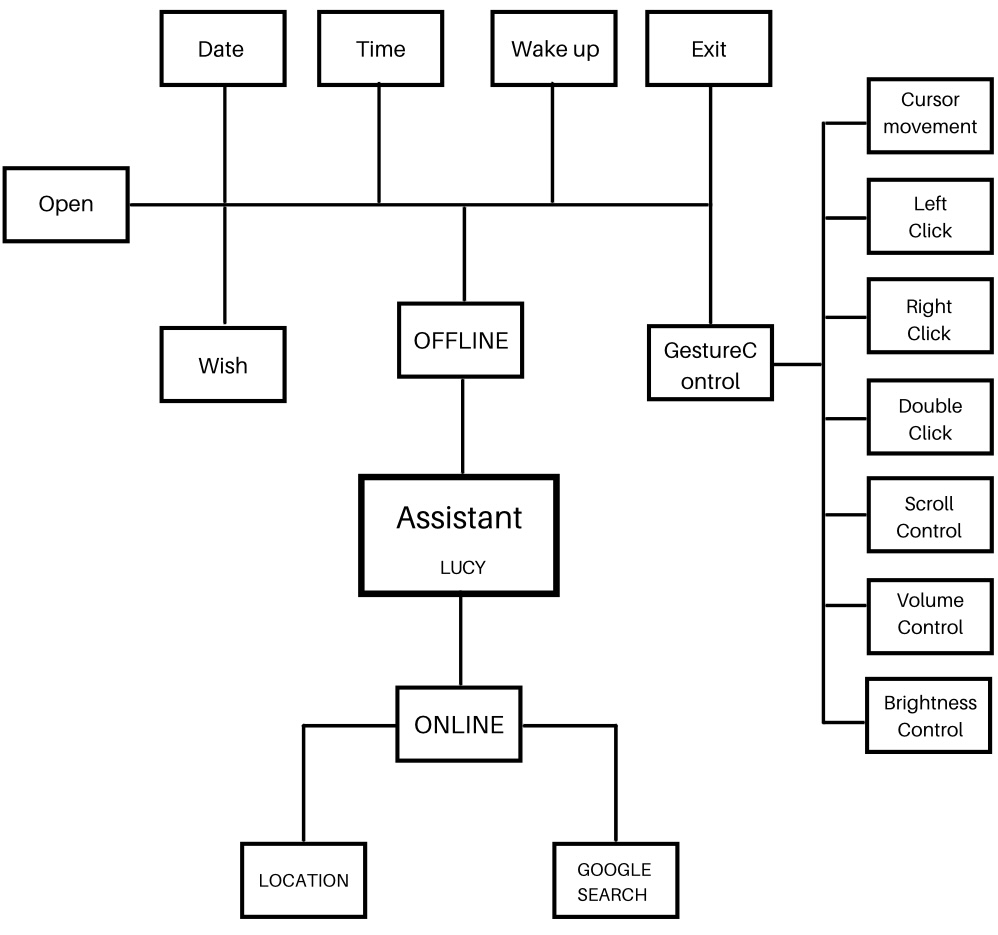


Figure 5.3.1 High Level Diagram

**CHAPTER 6**

**FUTURE PLAN**

**6.1 Future Plan**

* Use video feed from other high-quality cameras to improve efficiency.
* Static movements can help in sign language recognition.
* Add more voice commands, such as sending email, opening installed software, playing music/movies, and so on.
* Implement cloud-based software for home automation
* Include not only the ability to perform 3D gestures, but also the ability to work with multiple cameras to obtain a full 3D environment and achieve view-independent recognition, thereby eliminating some of the current system's limitations.

**CHAPTER 7**

**CONCLUSION**

The use of Digital Assistants is spreading at an alarming rate. Digital assistant technology is more than affordable and offers numerous advantages to its users. Automating repetitive tasks with a voice-activated personal assistant frees up human time and resources. Furthermore, it is capable of performing these mundane tasks efficiently and without error, which frequently results in increased customer satisfaction. While digital assistants handle routine tasks, humans can devote more time to responsibilities that necessitate human intervention for successful business solutions and services.

Controlling things with your hands and voice is more natural, easier, more flexible, and less expensive, and it has a lower failure rate. To encounter common issues and achieve a reliable result, it is obvious that much effort must be expended in developing reliable and robust algorithms with the assistance of a camera sensor with a specific characteristic.

Each of the techniques mentioned above has advantages and disadvantages, and they may perform well in some challenges while failing in others.

**CHAPTER 8**

**REFERENCES**

1) OpenCV Website – [www.opencv.org](http://www.opencv.org/)

2) MediaPipe Library - https://pypi.org/project/mediapipe/

3) H.S. Grif, C.C. Farcas, „Mouse Cursor Control System Based on Hand Gesture”, Procedia Technology, 22 (2016), 657-661

4) Lamberti, L.; Camastra, F. Real-time hand gesture recognition using a color glove. In Proceedings of the International Conference on Image Analysis and Processing, Ravenna, Italy, 14–16 September 2011; pp. 365–373. 10. Wachs, J.P.; Köl

5) O. Ozturk, A. Aksac, T. Ozyer, R. Alhajj, Boosting real-time recognition of hand posture and gesture for virtual mouse operations with segmentation. Applied Intelligence. 43(4) (2015) 786–801.