iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** X **Month of publication:** October 2022

DOI: https://doi.org/10.22214/ijraset.2022.47077

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue X Oct 2022- Available at www.ijraset.com

Gesture Controlled Mouse and Voice Assistant

Lokesh Gagnani[1], Hetvi Patel[2], Sakshi Chaturvedi[3], Rajeshwari Jaiswal[4]
[1, 2, 3, 4]*Computer Engineering Department, LDRP Institute of Technology and Research Gandhinagar Gujarat India*

Abstract: ***This project advocates a Human Computer Interaction method where cursor movement can be controlled via real-time camera by making use of human hand postures recognition. This method is an alternative to current techniques, which include manually pressing buttons or using physical computer mouse. Instead, it uses camera and computer vision software to manage different mouse events and can carry out any action that a conventional computer mouse can. The Virtual Mouse color recognition application will continuously gather real-time photos, which will then go through several conversions and filters. Once everything is converted, the application will use image processing to extract coordinates of specified color position from converted frames. The process then compares current color schemes within the frames to a list of color combinations, where each combination corresponds to a particular set of mouse operations. If the current color scheme matches, the application will perform the mouse command, which will be converted into a real mouse command on the user's computer. In addition to that authors have also developed a voice assistant to improve user productivity by managing routine tasks of the user and by providing information from online sources to the user because voice assistant is effortless to use.***
Keywords: Virtual Mouse, Hand Gestures, Media Pipe, Gesture Recognition, Voice Assistant, Machine Learning

I. INTRODUCTION

This project suggests a method for moving the cursor with just your hands, without the aid of any gadgets. The proposed system is designed to perform operations like right and left click, double click, drag and drop, multiple item selection, volume control etc. For the suggested system, the only necessary input device is a webcam. Python and OpenCV are the two pieces of software needed to put the suggested method into practice. On the system's screen, the camera's output will be shown so that the user may adjust it further. NumPy, math, wx, and mouse are the python dependencies that will be utilised to create this system.

The other functionality offered by this project is voice assistant. Voice searches have dominated over text search. The analysts are already predicting that 50% of searches will be via voice by 2024.Voice assistants are turning out to be smarter than ever. Voice Assistant is helpful to ease our day-to-day tasks such as showing date and time, performing searches on google, finding any location on google maps, opening any application etc. The voice assistants can take commands via text or by voice. An activating word, also known as a wake word, is required for voice-based intelligent assistants before the order may be given.For this project the wake word is quantum.

II. LITERATURE SURVEY

1) Chaithanya C, Lisho Thomas, Naveen Wilson, and Abhilash SS in 2018 proposed "Virtual Mouse Using Hand Gesture" where the model detection is based on colors.

2) AbhayDekate, ChaitanyaKulkarni, Rohan Killedar, "Study of Voice Controlled Personal Assistant Device", International Journal of Computer Trends and Technology (IJCTT) – Volume 42 Number 1 – December 2016.

3) Vinay Kr. Pasi, Saurabh Singh, and PoojaKumari in 2016 proposed "Cursor Control using Hand Gestures" in the IJCA Journal. The method suggests using various bands to carry out various mouse actions. The drawback is that different colors are required to carry out mouse tasks

4) Dung-HuaLiou, ChenChiung Hsieh, and David Lee in 2010. A real-time hand gesture recognition system using motion history images was proposed as a study. More intricate hand movements are this model's principal drawback.

III.BACKGROUND

A. Scope

Our objective is to offer more gestures so that users may complete more tasks quickly in the future. This proposal suggests a system that only makes use of the proper hand when making gestures. As a result, future improvements to the technique currently in use will allow for the use of both hands for various gestures. Many applications have been substantially improved through the rapid development of hand gesture recognition systems.

©IJRASET: All Rights are Reserved

Projects should accomplish the goal for which they were created, demonstrating the effectiveness of their execution. And through the voice assistant feature, authors have automated tasks such as finding locations on Google Maps, navigating documents, launching and stopping gesture recognition, performing Google searches, and sleeping / waking up the voice assistant. this function will benefit the user by saving them time and effort, as well as making computers more accessible to people who are blind or disabled.

B. The Root of Problem being Studied

The proposed Gesture controlled mouse system can be used to overcome problems in the real world such as situations where there is no space to use a physical mouse additionally for those who have hand issues and are unable to use a physical mouse. Moreover, there are various other problems like Physical mouse being subjected to mechanical wear and tear. It requires special hardware and a surface to operate. It is not easily adaptable to different environments and its performance varies depending on the environment.

Even in the current operational contexts, the mouse's capabilities are limited. All wired mouse and wireless mouse have its lifespan so the proposed mouse can be used to overcome these problems since hand gesture and hand Tip detection is utilized to operate the PC mouse using a webcam or an internal camera. Similarly, voice assistants could prove to be helpful for solving the problems related to the speed and errors that are faced when the searches are made by typing. It is a proven fact that voice is reputed to be four times faster than a written search: whereas one can write about 40 words per minute, one can speak approximately 150 words per minute.

IV. PROPOSED ALGORITHM

A. Gesture Controlled Mouse

1) *Step-1:* Initialization of Camera: The camera is utilized through an image as the user's finger moves in the camera window, it is essential to correlate the resolution of the system screen with the resolution of the camera window being used. So, it becomes important to determine how many pixels are moved on the actual screen if the user moves by 1 pixel in the camera window. This is necessary so that the user's finger can completely cover the screen while being photographed by the camera. Using the commands, the resolution of the screen can be estimated.
2) *Step-2:* Mapping and Extraction of all Hand Gestures: All the different gestures which are going to be used acquisition toolkit in order to capture the user's moving finger. A minimum of 15 frames per second must be the camera's capture rate.
3) *Step-3:* Co-relating the Resolution of Camera Window and Screen: Because the cursor will move across the entire screen are allocated binary numbers for mapping. Mediapipe package is used for extraction of different landmarks of the hand which is used to recognize gestures. Hand detection is done by the subtraction method which is easy and effective to detect the hand region from the original image. The binary numbers given are useful in fingers and palm segmentation. The palm point is defined as the center point of the palm with the help of landmark detection by the Mediapipe module and also it can be found by another method known as distance transform. Each pixel in the distance transform image keeps track of how far away the nearest border pixel is from it.

Binary values given to different gestures are:

FIST = 0
PINKY = 1
RING = 2
MID = 4
LAST3 = 7
INDEX = 8
FIRST2 = 12
LAST4 = 15
THUMB = 16
PALM = 31
V_GEST = 33
TWO_FINGER_CLOSED = 3
PINCH_MAJOR = 35
PINCH_MINOR = 36

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

4) *Step-4:* Recognition and Execution of Hand Gestures: Finally, after successful mapping and extraction of different landmarks of hand meaningful hand gestures are recognized and executed. This is done with the help of both OpenCV and Mediapipe modules. Depending on how many and what kinds of fingers are found, hand gestures can be anticipated. What fingers are detected depends on the content of the fingers? For example, if five fingers are detected it is classified as a neutral gesture, if two fingers are detected it is classified as a move cursor gesture, if the right finger is pulled down it is classified as a right click, etc.

B. Voice Assistant

The work started with analyzing the audio commands given by the user through the microphone. This can be anything like getting any information, operating the computer's internal files, etc.

The following algorithm shows how the voice assistant works.

1) *Step-1:* Speech Recognition: The system uses Google's online speech recognition system for converting speech input to text. The user's speech input is temporarily kept in the corpora organized on the computer network server at the information center so that they can access it. speech recognition system, which is subsequently transferred to Google cloud. The central processor is subsequently given the equivalent text, which is then received.
2) *Step-2:* Python Backend: The python backend gets the output from the speech recognition module and then identifies whether the command or the speech output is an API Call, Context Extraction, and System Call. The output is then sent back to the Python backend so that the user can receive the necessary output.
3) *Step-3:* Context Extraction: It is the task of automatically extracting structured information from unstructured and/or semi-structured machine-readable documents. This activity mostly involves using natural language processing to process documents written in human languages. Recent developments in the processing of multimedia documents, such as automatic annotation and content extraction from photos, audio, and video, can be considered context extraction test results.
4) *Step-4:* System Calls: In computing, a system call is a programmatic way in which a computer program requests a service from the kernel of the operating system it is executed on. This could involve interaction with hardware-related services (such as accessing a hard drive), starting and running new processes, and communicating with crucial kernel services like process scheduling. A crucial interface between a process and the operating system is provided via system calls.
5) *Step-5:* Text-To-Speech: Text-to-Speech (TTS) refers to the ability of computers to read text aloud. Written text is converted by a TTS Engine into a phonemic representation, which is subsequently translated into waveforms that can be output as sound. Written text is converted by a TTS Engine into a phonemic representation, which is subsequently translated into waveforms that can be output as sound.

V. IMPLEMENTATION

1) *Camera:* Using the Python computer vision library OpenCV, the video capture object is created, and the webcam will begin to capture video. The webcam then captures the frames and passes the ones to the AI virtual machine.
2) *Video Capture:* Detecting Which Finger is up and our job at this point is to detect which finger movements and the respective coordinates of the fingers are up based on the information found in the library and perform the mouse function accordingly. Keeping track of which finger is up based on the tip Id authors discovered via the MediaPipe and its associated coordinates, as well as carrying out the necessary actions, are required at this point. mouse function accordingly. Every 0.1 seconds, a loop is established to track the location of our hands.
3) *Hand Movement:* Mouse Features Using Computer Vision to Determine Hand Gestures and Hand Tip Detection Cursor Moving: The mouse pointer is made to travel across the computer's window using while the index and middle fingers are up for the Moving Function, use the Python AutoPy module.
4) *Left Button Click:* Left button click is made possible by wide opening the index finger and dragging it onto the file or folder.
5) *Right Button Click:* Wide opening the middle finger and dragging it onto the file or folder to explore more activities on it enables the right button to click.
6) *Scroll Up/Down:* The computer is designed to accomplish the mouse operation of scrolling up and down by merely pinching, which is done by connecting the tips of the index and middle fingers to the screen's scroll bar.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

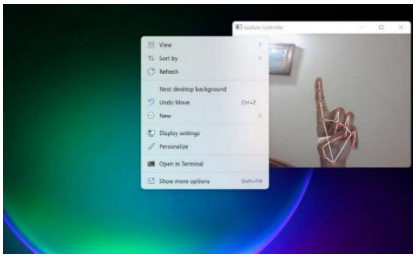


Fig. 1 Right Click



Fig. 2 Moving Cursor

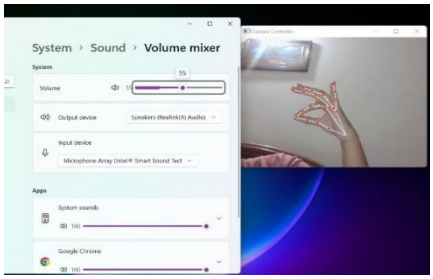


Fig. 3 Volume Adjustment

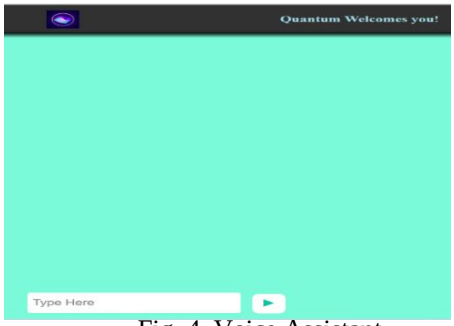


Fig. 4 Voice Assistant

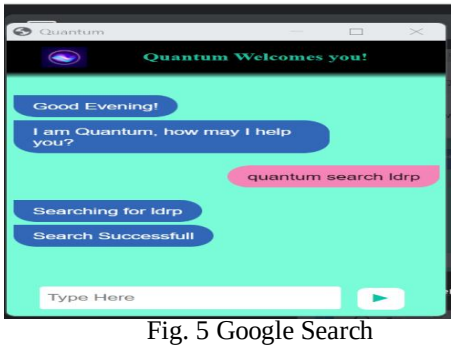


Fig. 5 Google Search

VI. CONCLUSIONS

In this paper, authors have discussed an alternative to the conventional physical mouse that provides mouse functions with the help of computer vision that houses a web camera that recognizes fingers and hand gestures and processes the captured frames and uses a machine learning algorithm to execute the defined mouse functions like moving the cursor, right-click, left click and scrolling function. After testing authors have concluded that the proposed virtual mouse system has worked exceedingly well and with great accuracy and the current system has overcome the drawbacks of the other systems. Through the Voice Assistant, authors have automated various services using a single-line command. It eases most of the tasks of the user like searching the web, file navigation, finding a location on google maps, opening any application, etc. The future plans include integrating quantum with mobile using React Native to provide a synchronized experience between the two connected devices. It is basically designed to minimize human efforts and control the device with just a human Voice.

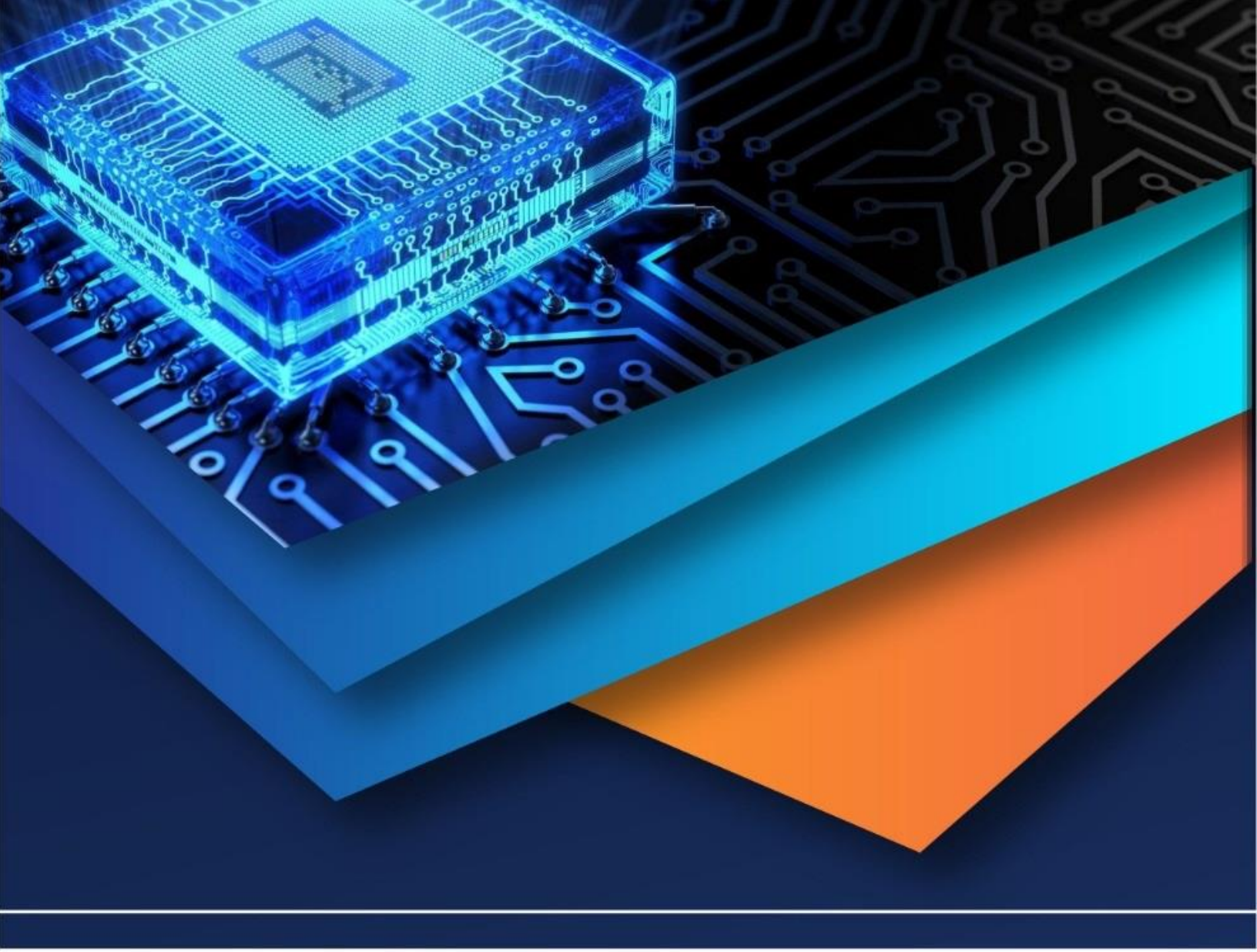
VII. ACKNOWLEDGMENT

This research paper would not have been possible without the guidance, assistance, and suggestions of many individuals. This acknowledgment transcends the authenticity of convention while we might express appreciation to all those persons on the back of the screen who directed and **encouraged** us for the accomplishment of our project. We would like to express our deepest appreciation and wholehearted thanks to Lokesh Gagnani. He has been a perpetual source of direction throughout the course of this research paper

REFERENCES

[1] https://www.hindawi.com/journals/tswj/2014/267872/
[2] https://courses.ece.cornell.edu/ece5990/ECE5725_Spring2019_Projects/Wednesday_051519/pa394_md848/index.html
[3] http://eprints.utar.edu.my/2262/1/IA-2016-13ACB07377-1.pdf
[4] https://www.irjet.net/archives/V5/i4/IRJET-V5I4872.pdf
[5] https://www.researchgate.net/publication/280112512_Hand_Gesture_Recognition_and_Cursor_Control
[6] https://www.studocu.com/in/document/lords-universal-college/human-machine-interaction/545737271-voice-assistant-project-report/282760

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)