Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, No.13, May : 2024

HAND GESTURE MOUSE CONTROLLER INTEGRATED WITH VOICE AUTOMATED CHAT BOT FOR DEVICEINTERACTION

Dr Ajay Talele, Professor, Dept. Of Computer Engineering, Vishwakarma Institute of Technology, Pune**.**
Pranay Junghare, Prithviraj Jadhav, Pratik Jadhav, Urmila Kakarwal, Student, Dept of Computer Engineering, Vishwakarma Institute of Technology, Pune.

Abstract
Human-computer interaction (HCI) is a field that is constantly evolving with the goal of creating digital interfaces that are easy for people to use. By offering a fresh and creative method for people to engage with computers, this research presents a novel approach to HCI. By combining voice and gesture controls, we're changing how users interact with technology. We're using advanced speech recognition technology and accurate hand gesture tracking to make this possible. Our system is different from traditional ways of interacting with computers because it embraces natural movements and expressions. This means users can control their computers simply by speaking and making hand movements. By combining these two ways of interacting, our system makes using computers easier and adds a personal touch. Additionally, we've added an AI chatbot to make interactions more dynamic, allowing users to have natural and friendly conversations with their computers. Through thorough studies, we've confirmed that our system works well and is easy to use. Beyond its immediate benefits, our research explores other ways this technology can be used, like in virtual reality or to help people with disabilities. By improving HCI methods, our work contributes to the ongoing discussion about how to make interactions with technology more engaging and user friendly.

Keywords:
Gesture Control, Human-Computer Interaction, Computer Vision, Media Pipe, ML, VoiceCommands, Hand Gestures

I. Introduction

The constant improvement of how people interact with computers is motivated by the goal of making digital interfaces easy to use and understand for users. In this context, our research takes a groundbreaking step, introducing a new and revolutionary way for people to interact with computers by combining voice and gesture controls. By using advanced speech recognition technologies and accurate hand gesture tracking made possible by Media Pipe, our system aims to change the way users interact with their computing devices. Traditional HCI methods often use input mechanisms that may be awkward or not naturally aligned with how humans naturally express themselves. Voice recognition, as a rapidly advancing field, adds a sound aspect to user-computer communication, allowing users to express their commands in a natural and clear way. At the same time, the involvement of gesture controls adds a physical aspect to this interaction, mimicking real-world movements and gestures. By combining these different ways of interacting, our system not only makes the user experience easier but also adds a touch of personalization, adjusting to each person's preferences and ways of interacting. Furthermore, the incorporation of an AI chatbot introduces a conversational element, expanding the system's capabilities beyond mere command execution. Users can have natural conversations, creating a mutually beneficial relationship between the user and the machine. This integration not only makes the system more useful but also adds to the ongoing conversation about the possible teamwork between artificial intelligence and human-computer interaction. As our research progresses, we thoroughly assess the system through user studies to measure its effectiveness, user satisfaction, and overall ease of use. Beyond the immediate advantages of a simpler and more user-friendly computing experience, our project looks into the wider consequences of this integrated

approach. Promising applications include serving virtual reality environments, providing features for users with different abilities, and creating useful productivity tools. Through this project, we aim to contribute to the progress of human-computer interaction, offering insights into how new technologies can work together and enhancing interactions between users and computers in various situations.

II. Literature

The research paper describes a system which employs Computer Vision, utilizing a webcam and HSV color detection in order to capture gestures of hand for cursor control. Users navigate and perform mouse operations with colored caps or tapes on their hands. Implemented in Python withthe OpenCV library, the real-time camera output is displayed on the monitor [1].

The research paper discusses emotion recognition which involves identifying human emotions, often through facial expressions and hand gestures. The study emphasizes that informal hand gestures, without a formal system, can deliver meanings independently of facial expressions.

Additionally, it explore how expressions of face and

casual hand gestures can synergize to create entirely new messages. With the use of Convolutional Neural Networks with TensorFlow and OpenCV, they developed an image classification algorithmthat recognizes emotions conveyed solely through hand gestures [2].

The article examines six attack surfaces that encompass the gathering, sending, processing, and storing of voice data within the Alexa ecosystem. Concerns regarding attacks on both frontend voice capture and backend voice command recognition are addressed, along with proposedmitigation strategies to improve security and privacy for Alexa and other voice assistants. [3].

This research study develops a voice chatbot device to alleviate the difficulties visually impaired people encounter when transacting financial business. Using AI technologies, the chatbot uses Amazon Lex and Eyowo as the payment platform to streamline transactions. [4].

This research addresses a software application which mimics human conversation, aiming to reduce the time users spend searching for information. The focus here is on a keyword recognitionchatbot, which interprets user queries and promptly delivers relevant answers from its database [5].This research employs design science analysis to enhance chatbot design awareness in higher education. A literature review and analysis of student communications in a Java programming course inform the conceptual architecture for educational chatbots [6].

The main goal of this project is to employ AI and machine learning to build an API and chatbot for Matrusri Engineering College. As users interact with the chatbot, it changes and becomes moreaccurate over time. The chatbot's goal is to improve the user experience on the college website byusing machine learning and natural language processing to provide information and support through natural language input. [7].

This study presents PYNQ_Z2 and gesture recognition as the foundation for an intelligent ordering system. For gesture segmentation, the algorithm uses skin detection, background subtraction, andface removal in the YCbCr color space [8].

The paper introduces Cognitive Info communications (Cog Infocom), a new field that combines ICT with cognitive science. It explores how human thinking can be mixed with digital devices forsmoother interactions. It talks about the latest progress, trends, and new ideas in Cog Infocom, especially how it's making learning and using computers better [9].

The research paper explores various methodologies and systems developed for hand gesture recognition in man-machine interfaces. It covers research using different camera technologies, such as single lens, multi-lens, depth perception lens, or infra-red lens. Methodologies which are discussed include principle component analysis, hidden Markov model, particular filtering, finite state machine, and neural networks. The survey also examines specific applications like media player control, robot interaction, and human-computer interaction, along with their limitations andadvancements [10].

The research discusses the emergence of virtual mouse technology as a response to problems facedby

customers when purchasing traditional mice online, especially during the pandemic. It highlights the shift towards contactless interaction and the utilization of webcams for hand gesture recognition, eliminating the need for additional hardware. Additionally, the review emphasizes thepredominant use of Artificial Intelligence and Machine Learning techniques in developing such systems, emphasizing their significance in advancing Human-Computer Interaction [11].

The paper introduces a new system where you can control the computer cursor using hand movements captured by a webcam. It highlights how hand gestures are easy for everyone to understand and use, suggesting a system that lets you control the mouse and transfer files betweencomputers on the same network. This system works by preparing the hand movements, identifyingkey features, and matching them to predefined gestures, making it easier for people to interact withcomputers naturally [12].

III. Components and Modules Used

1. OpenCV

It's like a toolbox for computers to observe or see and understand images and videos. It helps in building smart computer vision applications like robots that can navigate, systems that recognize faces and gestures, and tools for analysing features in 2D and 3D images.

2. Media Pipe

It is a versatile framework created by Google for building smart systems that can understand different types of information, like tracking movements, detecting objects, recognizing faces, and even understanding hand gestures. It's handy for developers working on various types of projects.

3. NumPy

NumPy is like a magic wand for dealing with numbers and arrays (kind of like lists, but more powerful) in Python. It's super-fast and makes it easy to do complex mathematical operations, like Fourier transforms, random number generation, and even linear algebra.

4. Pycaw

This is a special library for working with audio on Windows computers. It helps with tasks like removing background noise from audio recordings.

5. TKinter

It is used to create windows and buttons and other graphical elements in your Python programs, TKinter is a go-to tool. It's a simple and easy way to build user interfaces for your applications, and it works on most computers.

6. PyAutoGUI

To control your computer just by writing a Python script PyAutoGUI is used. It is used to move the mouse around, click on things, type text, take screenshots, and even send messages and alerts to yourself. It's like having a robot assistant for your computer tasks.

7. Webcam As Hardware

The only hardware component used is the webcam which serves as the primary input device, capturing live video feed of hand gestures.

IV. Methodology

The project introduces the Gesture Controlled Virtual Mouse and Voice controlled Chatbot, a system that transforms human-computer interaction by combining Hand Gestures and Voice Commands, reducing the need for direct physical contact in all input/output operations. The hardware setup includes standard components (for the project we have used the camera of the pc or laptop) with a focus on hand gesture recognition, eliminating the need for additional hardware.The software setup, using Python version 3.8.5, includes advanced Machine Learning and Computer Vision algorithms, particularly a CNN implemented by Media Pipe through pybind11. Gesture recognition involves using the latest and most advanced algorithms for accurately identifying both still and moving hand gestures. Recognized gestures cover a broad range,including Neutral Hand Gesture, Move Cursor, Mouse Left

Click, Mouse Right Click, Double LeftClick, Scroll, Dragging and Dropping, Multiple Item Selection, Controlling of Volume of the device, and controlling the brightness of the device. The Voice Assistant, named Proton, brings a conversational aspect, allowing commands like starting/stopping Gesture Recognition, Google Search, finding locations on Google Maps, navigating files, retrieving the current date and time, copying and pasting, putting Proton to sleep or waking it up, and exiting the system. The project uses Machine Learning models, specifically CNN (convolutional neural network) implemented byMedia Pipe running on top of pybind11, to ensure accurate gesture recognition of hand. The implementation is customized specifically for the Windows platform.

The actions to be performed using gesture recognition are as follows:-

1. Neutral Gesture (It refers to a specific hand position that serves as a baseline or starting pointfor recognizing and interpreting other gestures).
2. Moving the mouse cursor
3. Mouse Left Click
4. Mouse Right Click
5. Double Left Click
6. Scrolling from top to bottom and vice versa
7. Dragging and Dropping
8. Multiple Item Selection
9. Controlling the Volume
10. Brightness Increment or decrement

The Chatbot works on limited commands which are as follows:-

1. Launch / Stop Gesture Recognition
2. To browse anything on Google
3. Find a Location using Google Maps
4. To open any file
5. Display Current date and time
6. Copy and Paste
7. To pause the voice command feature and to wake it up.
8. Exit the chatbot

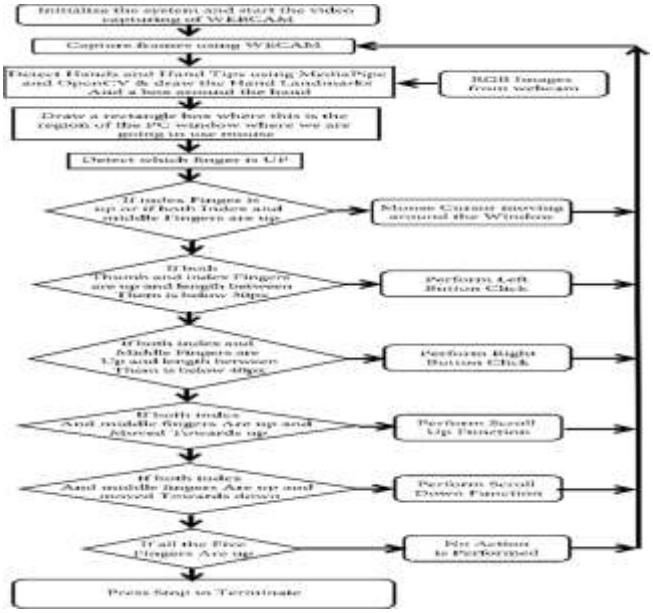

Figure 1: Flowchart of working of the Gesture Recognition Module and functioning according tothe fingers used.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, No.13, May : 2024

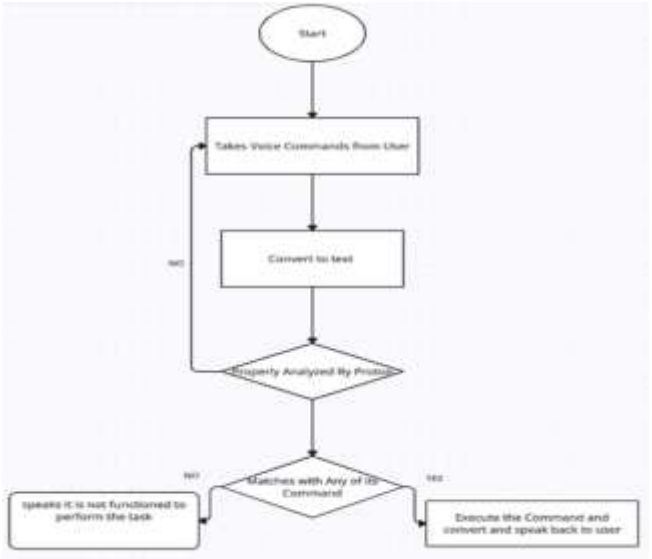

Fig. 2. Flowchart of working of the AI Voice Chatbot

V. Results and Discussion

The Gesture Controlled Virtual Mouse project signify a notable achievement in redefining human-computer interaction. The testing results were satisfactory as the system seamlessly combined hand gestures and voice commands, making it easy and natural to use. Appreciation was received as the system could be used by directly moving your hands. To analyze how well the system worked, weused some measurements like how fast it responded and how accurate it was. The results indicatedthat the system rapidly identified hand gestures and responded quickly to voice commands. This was possible due to the algorithms we used, especially the CNN by Media Pipe, which accuratelyidentified a variety of hand gestures. The following pics shows the results of the project.

Fig.3. Neutral Base Gesture & Moving Cursor Gesture

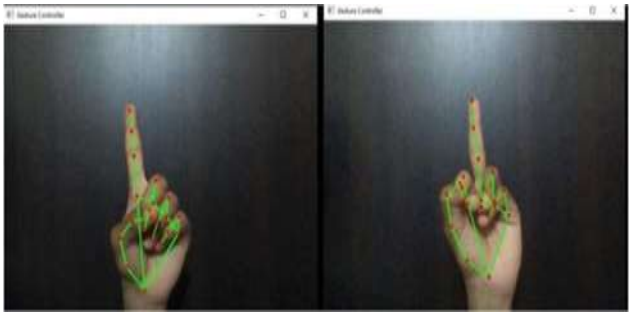

Fig.4. Mouse Left Click & Mouse Right Click

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, No.13, May : 2024

Fig.5. Double Left Click & Dragging and Dropping

Fig.6. Controlling the Brightness

Fig. 7. The figure displays the normal conversation with the AI chatbot.

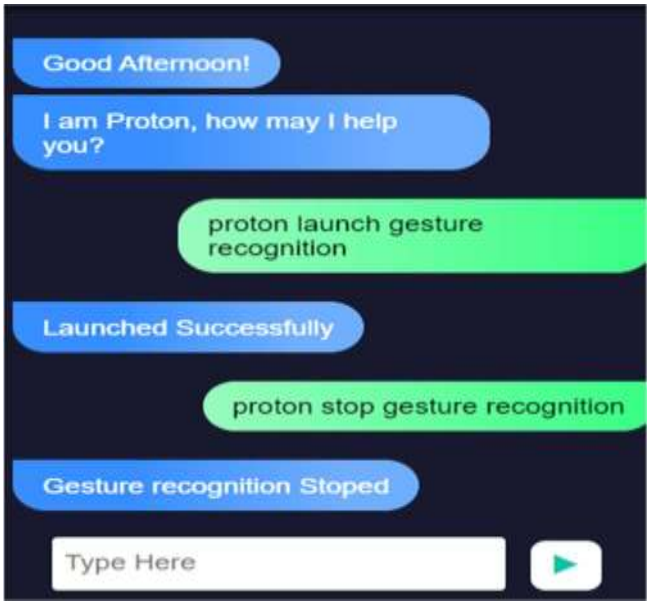

Fig. 8. This figure displays the launch and stop command of the AI chatbot.

In our discussion, we highlight how the project is important for Human-Computer Interaction. We highlight how the system's flexibility, including features like neutral gestures, moving the cursor, clicking, scrolling, and voice commands by Proton, makes it suitable for various needs. The glove-based interaction is not just another way to control the system; it also makes it easier for users to access and use the system.

We also discuss the future potential of the project. It creates opportunities for further research in areas such as making technology easier to use, virtual reality experiences, and tools that make workmore efficient. In summary, our project not only achieves its goals of providing a fun way to use computers but also adds to the ongoing discussions about how people and computers interact.

VI. Conclusion

In conclusion, our integrated voice and gesture-controlled system transforms user-computer interaction, providing a smooth and personalized computing experience. Through the combinationof voice commands, accurate hand gestures, and an AI chatbot, we have created a dynamic interface that goes beyond traditional ways of doing things. The successful combination of these methods not only improves efficiency but also creates opportunities for innovative applications. As technology advances, our work helps in the continuous development of HCI, paving the way for a future in human-computer interactions that feels more natural, easy to understand, and deeplyengaging.

References

[1] M. Shetty, C. A. Daniel, M. K. Bhatkar and O. P. Lopes, "Virtual Mouse Using Object Tracking," 20205th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India,2020, pp. 548-553, doi: 10.1109/ICCES48766.2020.9137854

[2] A. M. Arjun, S. Sreehari and R. Nandakumar, "The Interplay Of Hand Gestures And Facial Expressions In Conveying Emotions A CNN–BASED APPROACH," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2020, pp. 833-837 doi:10.1109/ICCMC48092.2020.ICCMC-000154.

[3] Y. Lit, S. Kim and E. Sy, "A Survey on Amazon Alexa Attack Surfaces," 2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 2021, pp. 1-7,doi: 10.1109/CCNC49032.2021.9369553.

[4] I. Samuel, F. A. Ogunkeye, A. Olajube and A. Awelewa, "Development of a Voice Chatbot for Payment Using Amazon Lex Service with Eyowo as the Payment Platform," 2020 International Conference on Decision Aid Sciences and Application (DASA), Sakheer, Bahrain, 2020, pp. 104-108, doi: 10.1109/DASA51403.2020.9317214.

[5] M. D. Thakkar, C. U. Sanghavi, M. N. Shah and N. Jain, "Infini – A Keyword Recognition Chatbot," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, 2021, pp. 1036-1042, doi: 10.1109/ICAIS50930.2021.9395818.

[6] M. Rakhra et al., "E-Commerce Assistance with a Smart Chatbot using Artificial Intelligence," 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom, 2021, pp. 144-148, doi: 10.1109/ICIEM51511.2021.9445316.

[7] H. K. K., A. K. Palakurthi, V. Putnala and A. Kumar K., "Smart College Chatbot using ML and Python," 2020 International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, 2020, pp. 1-5, doi: 10.1109/ICSCAN49426.2020.9262426.

[8] Z. Meng, K. Xie and L. Yang, "An intelligent ordering control system based on hand gesture recognition," 2020 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS), Shenzhen, China, 2020, pp. 1-5, doi: 10.1109/HPBDIS49115.2020.9130570.

[9] Katona, Jozsef. (2021). A Review of Human–Computer Interaction and Virtual Reality Research Fieldsin Cognitive Info Communications. Applied Sciences. 11. 2646. 10.3390/app11062646.

[10] D.-H. Liou, C. -C. Hsieh and D. Lee, "A real time hand gesture recognition system using motion

history image," 2010 2nd International Conference on Signal Processing Systems, Dalian, China, 2010, pp. V2-394-V2-398, doi: 10.1109/ICSPS.2010.5555462.
[11] K. P. Vinay, "Cursor control using hand gestures," International Journal of Critical Accounting,vol. 0975–8887, 2016.
[12] L. Thomas, "Virtual mouse using hand gesture," International Research Journal of Engineeringand Technology (IRJET, vol. 5, no. 4, 2018).