Browser Control with Hand Gesture

Manoj kr. Yadav, Aditya Kumar, Prince Bhardwaj, Akash Gupta, Anshu kr.verma, Abhinav Choudhary
Department of Computer Science and Engineering, IMS Engineering College, Ghaziabad, India

Abstract -Speech recognition turns spoken words into text, while gesture recognition identifies human movements. With our real-time gesture detection system, users can guide the computer to specific websites using gestures in front of a webcam. We also have a voice recognition system that translates spoken words into text and performs tasks accordingly. Traditional input devices like mice and keyboards require manual handling. Visual and auditory cues are vital for one-way communication between humans and computers. Hands are commonly used for physical tasks and communication. Speech and gestures help us convey messages clearly in daily conversations. For mute and deaf individuals who rely on gestures and hands for communication, these tools are crucial.

INTRODUCTION

Hand motions are crucial in sign language for communication. If computers could interpret and translate hand gestures, it could greatly enhance human-computer interaction. This requires processing extensive information from images, which is essential for implementing such technology..

Yet, even with this capability, users may face limitations if they can't use their hands freely, such as elderly or disabled individuals. In such cases, using speech or voice to operate equipment or systems can be more efficient than relying on someone else.

Instead of typing on a keyboard, speech recognition technology allows users to input data by speaking into a microphone, providing the same results as manual typing. This technology has advanced significantly in recent years.

Gesture recognition research aims to develop systems that can interpret human movements to convey information or control devices. Various methods used for gesture recognition include hand detection segmentation, RGB color schemes, webcams, real-time tracking, and Markov hidden models. These techniques help computers understand and interpret human gestures., and depth map techniques, are used in this research.

To address these issues, I plan to create a system that uses image processing and hand gestures to control the browser. Users would input their commonly visited websites and navigate them using hand gestures. Specific hand movements would correspond to specific websites. With this system, users can control the browser, open websites in new tabs, and close the browser using hand gestures. This creates a hand gesture-controlled browser, allowing users to quickly access their chosen websites in new tabs.

Problem Statement

The rapid growth of the internet has increased the use of web browsers, with users frequently visiting several favorite websites. Simplifying access to these sites would enhance user experience. A potential solution is integrating hand gestures and speech recognition into web browsers. Given that web browsers are used constantly, this technology could make browsing more intuitive and efficient, providing users with an easier and more seamless way to control their web activities.

Motivation

To enhance user-friendliness and accessibility, we are developing a Chrome extension that enables users to perform various actions on websites using hand gestures. This extension works across all Chromium-based browsers and can potentially be extended to other browsers. The system operates entirely on the user's computer, ensuring privacy and security of user data. Local processing results in low latency for gesture recognition, making this system an efficient and secure way for users to interact with websites using hand gestures.

Objective

Gestures and speech are major forms of human communication. The primary goal of gesture recognition is to create a system that can identify specific human gestures, while the primary goal of speech recognition is to convert speech sounds into corresponding text. In this project, we have

implemented a real-time gesture recognition system that allows a user to navigate a computer to a particular website by performing a specific gesture in front of a web camera connected to the computer. Additionally, we have implemented a speech recognition system that takes speech input from the device's microphone, converts it into text, and then performs various operations based on the recognized text.The integration of gesture and speech recognition in web browsers offers a more natural and efficient way for users to interact with websites. By leveraging local processing, the system maintains user privacy and delivers real-time performance, making it a practical and secure solution for enhancing web navigation..

Implementation

The use of gesture-based interaction has been increasing, and vision-based devices may eventually replace the mouse and keyboard. The main advantage of using hand gestures with a computer is the natural interaction they provide. Gestures, which are nonverbal communication involving body movements, mostly originate from the hands and face. They offer a communication method that mimics real-world interactions and doesn't require additional technology to function.

There are two types of hand gestures: static and dynamic. Static gestures involve hand postures, while dynamic gestures involve movement over time. Data glove techniques, vision-based methods, and colored-marker approaches are used to gather data for hand motion recognition.

Vision-based techniques use cameras to capture human motion and can handle various aspects of gestures, such as color and texture, which sensors lack. However, they face challenges like diverse lighting, complex backgrounds, and objects that resemble hands.

Creating a robust and real-time gesture detection system involves overcoming obstacles like complex hand structures, different lighting conditions, real-time processing, and background noise. This study uses the running average concept in background subtraction to identify and extract the hand from the background. Human hands are tagged with colorful markers or gloves, aiding in hand tracking and finger and palm detection. Marker gloves can extract geometric characteristics to define the hand's contour. A wool glove with three different colored fingers is used to represent palms and fingers.

This technique is simpler and cheaper compared to sensors or data gloves, but it still lacks natural human-computer interaction.

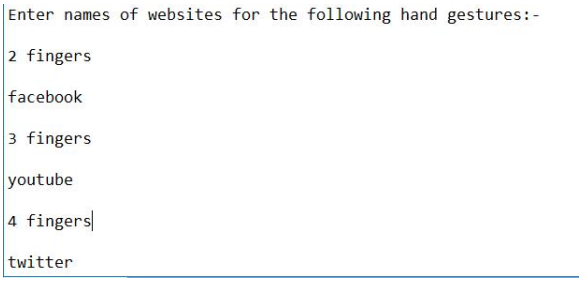
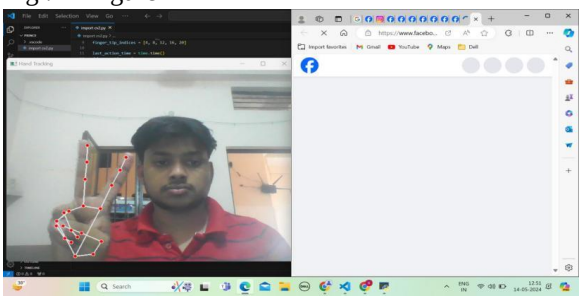


Fig .1-Figure

Fig .2 - Figure



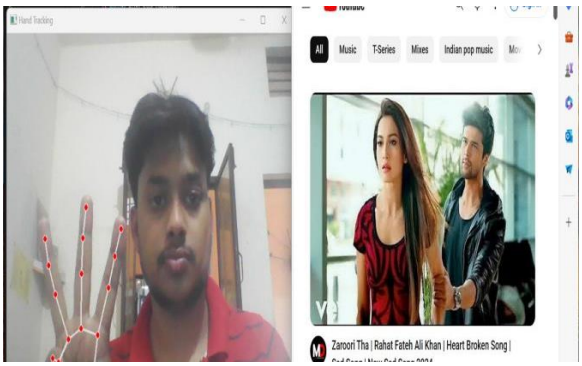


Fig .3 - Figure

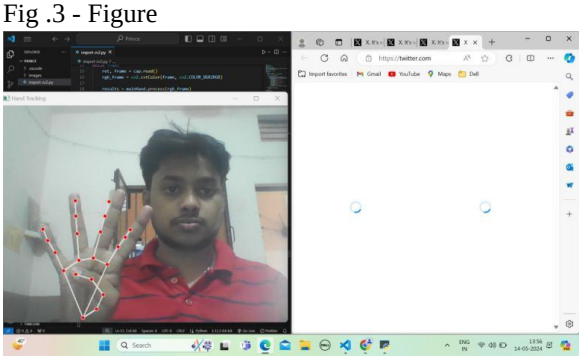


Fig .4 - Figure

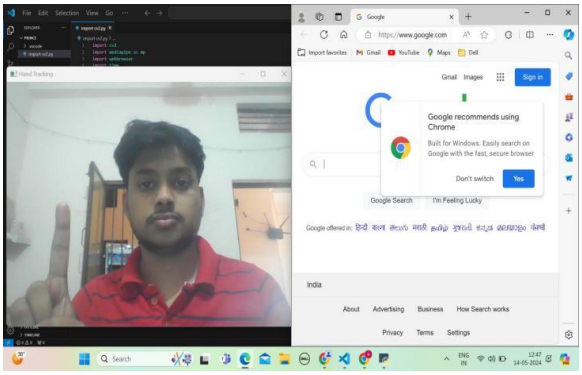


Fig .5 - Figure

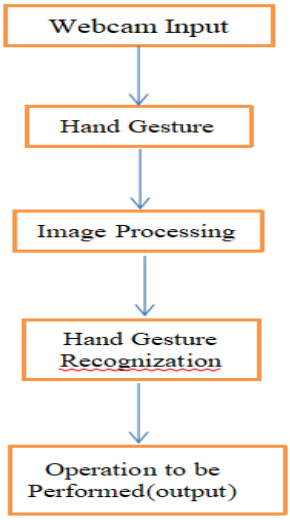


Fig .6 - Figure

Literature Survey

In the computer science and technology era, gesture recognition is a crucial field that translates human gestures using various computer vision techniques and algorithms. While numerous human body motions can create gestures, the most common are generated from the face and hands. The process of tracking gestures, interpreting them, and converting them into useful commands is known as gesture recognition. Various techniques and methods have been employed to design and develop systems capable of performing this task

[1].”M. KRUEGER” Artificial reality II, Addison-Wesley Reading (Ma)The starting approach of interaction with a computer using hand gestures was first projected by

Myron W. Krueger in 1970.

[2]”The paper titled "Media Control Using Hand Gestures with VLC Media Player" by Singh, M., and Rizvi, S., published in the Journal of Informatics Electrical and Electronics Engineering in March 2021, discusses the development of a gesture-based control system for managing media playback in VLC Media Player using hand gestures. This research integrates computer vision and human-computer interaction techniques to provide a more intuitive and natural way for users to interact with media playback software without the need for traditional input devices such as keyboards, mice, or remote controls.

[3]."Controlling Devices Using Hand Gesture" by Sampath, M., Velraj, P., Raghavendran, V., and Sumithra, M., published in the World Journal of Advanced Research and Reviews in June 2022, explores a hand gesture recognition system designed to control various electronic devices. This work represents a significant advancement in the field of human-computer interaction (HCI), providing a more natural and intuitive way for users to interact with technology using simple hand gestures.“J

[4]. "Controlling Computer using Hand Gestures" by Vijayakumar, D. and Brahanyaa, S., published in the Journal of Emerging Technologies and Innovative Research in April 2019, discusses the development and implementation of a gesture recognition system to control computer functions using hand gestures. The study integrates principles of computer vision and machine learning to create a user interface that interprets physical gestures as commands to perform specific tasks on a computer, such as opening applications, navigating web pages, or controlling multimedia playback.

[5]."Controlling web browser via different hand gestures and speech recognition" by Kedari, P., Kadam, S., and Prasad, R., published in the International Research Journal of Engineering and Technology in June 2022, explores a hybrid approach to enhancing user interaction with web browsers through the use of both hand gestures and speech recognition

[6]."Browser Control Using Hand Gestures" by Chapalgaonkar, V., Kulkarni, A., and Sonawale, A., published in the International Journal for Research in Applied Science & Engineering Technology in April 2022, discusses the development and implementation of a system for controlling media playback using hand gestures. The study integrates computer vision techniques to recognize hand gestures and translate them into commands to control media playback.

[7].“Pei Xu”,In this, only one hand moment is used to control the TV. A hand picture looks like an icon that follows the hand movements of the user appearing on the screen display on the TV. In this paper

[8]."Jadhav, R.V., Bait, P.S., and Pundkar, S.’’, and published in the International Journal of Advance Research, Ideas and Innovations in Technology in 2022, explores the development and implementation of a system for controlling computers using hand gestures. The study focuses on leveraging computer vision techniques to interpret hand movements as commands for various computer operations.

[9]“ Jae-Hun Song, Kyeongbo Kong, and Suk-Ju Kang”,The primary goal is to develop an accurate and efficient system for recognizing dynamic hand gestures from video sequences. Dynamic gestures involve movements over time, which adds complexity to the recognition task compared to static gestures.

[10]”Rohit Mukherjee, Pradeep Swethen, Ruman Pasha, Sachin Singh Rawat”noted that while the developed gesture recognition model is user-friendly compared to the latest available systems or command-based systems, it has a significant drawback: it is less powerful in accurately spotting and recognizing gestures. The current system struggles in complex environmental backgrounds and normal lighting conditions, indicating a need for further improvement and the development of a robust network for gesture recognition.

The existing model, designed for six gesture classes, can be used to control various applications such as PowerPoint presentations, Windows Picture Manager, media players, and games. However, to enhance its performance and reliability, particularly in diverse and challenging environments, further advancements in the system are required.

[11].”Rohit Mukherjee, Pradeep Swethen, Ruman Pasha, Sachin Singh Rawat”The hand gesture recognition system based on Arduino Uno and ultrasonic sensors offers an innovative way to manage devices like media players and web browsers. By using gestures to control VLC media player operations such as playing, pausing, and scrolling, the system provides a hands-free and interactive experience. This makes it especially suitable for educational environments, enhancing the teaching and learning experience through more engaging and interactive methods.Arduino Uno: A microcontroller board used to process input from the sensors and communicate with the laptop.

Ultrasonic Sensors: Devices that detect hand gestures by measuring the distance between the sensor and the user's hand.

Python: Used to create a serial connection between the Arduino and the laptop, allowing for the control of applications.

[12].”Ayushi Bhawal, Debaparna Dasgupta, Arka Ghosh, Koyena Mitra, Surajit Basak”,In their work, Ayushi Bhawal, Debaparna Dasgupta, Arka Ghosh, Koyena Mitra, and Surajit Basak describe a system that leverages Arduino Uno and ultrasonic sensors to control various operations on a laptop, such as media player functions and volume adjustments. The system uses Arduino, ultrasonic sensors, and Python to establish a serial connection for communication between the hardware and the laptop.Media Player Control: Hand gestures detected by the ultrasonic sensors can play, pause, and manage other functions of a media player such as VLC.

Volume Adjustment: Gestures can increase or decrease the volume of the media player.

Page Scrolling: Hand gestures can scroll up and down through pages on the laptop, making it easier to navigate documents and web pages without a keyboard or mouse.

[13]. ”Sarita K., Gavale Yogesh, S. Jadha”,The paper discusses the development of a hand gesture monitoring system that utilizes ultrasonic sensors and is built on the Arduino microcontroller ATMEGA32. This system stands out for its simplicity and cost-effectiveness, as it claims to require no additional hardware for classifying hand gestures beyond the Arduino and ultrasonic sensors.

The proposed hand gesture monitoring system offers an innovative approach to human-computer interaction through the use of basic hardware components. By leveraging the ATMEGA32 microcontroller and ultrasonic sensors, the system provides an efficient and cost-effective solution for gesture recognition. This simplicity ensures that the technology can be implemented in a variety of settings, from personal DIY projects to more sophisticated educational and commercial applications.

[14]”Chuankun Li, Shuai Li, Yanbo Gao, Xiang Zhang, and Wanqing Li”, published in the IEEE

Transactions on Cognitive and Developmental Systems, volume 14, issue 4, pages 876-886, in December 2022, introduces a novel approach to The research on hand gesture recognition using a two-stream neural network architecture based on pose estimation focuses on enhancing the accuracy and robustness of recognizing gestures by effectively capturing and analyzing both spatial and temporal information related to hand movements[15]The presented real-time on-device hand tracking pipeline is designed to predict hand skeletons from a single RGB camera, specifically for augmented reality (AR) and virtual reality (VR) applications. This pipeline is implemented using MediaPipe, a versatile framework for creating cross-platform machine learning solutionsPalm Detector: This initial model quickly identifies the presence and location of a palm within the camera frame. The palm detector is optimized for speed and accuracy, ensuring that the subsequent hand landmark model can focus on the relevant area of the image.

Hand Landmark Model: Once the palm is detected, this model takes over to predict the precise locations of the hand's skeletal keypoints. It maps the hand's structure by identifying various landmarks, such as joints and fingertips.

CONCLUSION

In conclusion, as technology continues to advance and hand gestures remain a staple in everyday communication for conveying clear intentions, hand gesture recognition has emerged as a crucial component of Human-Computer Interaction (HCI). This technology allows devices to detect and classify hand gestures, performing corresponding actions to make interactions more natural and straightforward.The integration of hand gestures with speech recognition to control a web browser represents a significant advancement in HCI. This technology provides users with a more immersive and accessible computing environment, simplifying interactions and making technology more user-friendly. By continuing to develop and refine these systems, we can pave the way for even more intuitive and effective methods of human-computer interaction, broadening the scope and impact of gesture recognition technology across various domains.

ACKNOWLEDGEMENT

To define the concept of a convex set, any line connecting two points within the set must lie entirely within the set. For hand gesture recognition, the convex hull is used to outline the area of the hand, defining the boundary that contains all other points of the hand within it. To locate the center of the palm, extreme points on this hull are identified. A circle is then drawn around the fingers, centered at the palm with a radius equal to 70% of the longest distance from the center to these extreme points. This setup helps distinguish fingers by intersecting a circle with a thresholded image of the hand, using bitwise operations to identify different fingers.. Gesture recognition is dynamic; after recognizing and processing a gesture, the system captures a new image to start the process again, continuously adapting to new gestures.

REFERENCES

[1] M. KRUEGER, Artificial reality II, Addison-Wesley Reading (Ma), 1991.

[2] Singh, M. and Rizvi, S., "Media Control Using Hand Gestures with VLC media player," Journal of Informatics Electrical and Electronics Engineering, Mar. 2021

[3] Sampath, M., Velraj, P., Raghavendran, V., and Sumithra, M., "Controlling Devices Using Hand Gesture," World Journal of Advanced Research and Reviews, Jun. 2022.

[4] Vijayakumar, D. and Brahanyaa, S., "Controlling Computer using Hand Gestures," Journal of Emerging Technologies and Innovative Research, Apr. 2019

[5] Kedari, P., Kadam, S., and Prasad, R., "Controlling web browser via different hand gestures and speech recognition," International Research Journal of Engineering and Technology, Jun. 2022.

[6] Chapalgaonkar, V., Kulkarni, A., and Sonawale, A., "Media Control Using Hand Gestures," International Journal for Research in Applied Science & Engineering Technology, Apr. 2022.

[7] Pei Xu, "A Real-time Hand Gesture Recognition and Human-Computer Interaction System", arXiv:1704.07296v1 [cs.CV] 24 Apr 2017.

[8] Jadhav, R.V., Bait, P.S., and Pundkar, S., "Controlling Computer using Hand Gestures," International Journal of Advance Research, Ideas and Innovations in Technology, 2022.
[9] Jae-Hun Song, Kyeongbo Kong, and Suk-Ju Kang. "Dynamic Hand Gesture Recognition Using Improved Spatio-Temporal Graph Convolutional Network." IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 9, September 2022, pp. 6227.
[10] Rohit Mukherjee, Pradeep Swethen, Ruman Pasha, Sachin Singh Rawat, "Hand Gestured Controlled laptop using Arduino", International Journal of Management, Technology And Engineering, October 2018.
[11] Ayushi Bhawal, Debaparna Dasgupta, Arka Ghosh, Koyena Mitra, Surajit Basak, "Arduino Based Hand Gesture Control of Computer", IJESC, Volume 10, Issue No.6, June 2020.
[12] Udit Kumar, Sanjana Kintali, Kolla Sai Latha, Asraf Ali, N. Suresh Kumar, "Hand Gesture Controlled Laptop Using Arduino", April 2020.
[13] Sarita K., Gavale Yogesh, S. Jadhav, "HAND GESTURE DETECTION USING ARDUINO AND PYTHON FOR SCREEN CONTROL", International Journal of Engineering Applied Sciences and Technology, 2020, Vol. 5, Issue 3, ISSN No. 2455-2143, July 2020.
[14] Chuankun Li, Shuai Li, Yanbo Gao, Xiang Zhang, and Wanqing Li, "A Two-Stream Neural Networkfor Pose-Based Hand Gesture Recognition," IEEE Transactions on Cognitive and Developmental Systems, vol. 14, no. 4, pp. 876-886, Dec. 2022.
[15] Fan Zhang, Valentin Bazarevsky, Andrey Vakunov, Andrei Tkachenka, George Sung, Chuo-Ling Chang, and Matthias Grundmann. "MediaPipe Hands: On-device Real-time Hand Tracking." arXiv preprint arXiv:2006.10214 (2020).