International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Conference Proceedings: Volume: 05 ACME – 2026 DOI: 10.55041/ISJEM.ACME095
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Hand Gesture Recognition using HCI

Mrs P Sujitha [1]
Assistant Professor, Department of AI&DS
Annamacharya Institute of Technology and Sciences, Tirupati – 517520, A.P.
sujithapannapalli@gmail.com

C Varun Kumar Reddy [2]
Department of AI&DS
Annamacharya Institute of Technology and Sciences, Tirupati – 517520, A.P.
varunreddy7456@gmail.com

N Sriram [3]
Department of AI&DS
Annamacharya Institute of Technology and Sciences, Tirupati – 517520, A.P.
ramnsriram790@gmail.com

Shabreen Shaik [4]
Department of AI&DS
Annamacharya Institute of Technology and Sciences, Tirupati – 517520, A.P.
shaikshabreen919@gmail.com

C Nandini [5]
Department of AI&DS
Annamacharya Institute of Technology and Sciences, Tirupati – 517520, A.P.
cnandini8654@gmail.com

ABSTRACT — There is an increased popularity in touchless and more natural interfaces to interact with computers. The following paper suggests a virtual mouse that is based on the vision principle and is controlled by using the movement of hands that have been captured by a regular web camera. Hands are real-time detected and translated into computer actions. The movement of the cursor, left and right-clicks, double-clicks, scrolling, switching, volume and brightness of the screen are some of the features of this system.

The system works by first capturing video frames, detecting the landmarks of the hands, recognizing the gestures, and then mapping the detected gestures to different actions. This is done without the need for additional hardware, thus providing a more practical alternative for interacting with the computer. Real-time results are obtained, proving the feasibility of the system. This system is an improvement over the existing interfaces for interacting with the computer.

Keywords — Human–Computer Interaction, Virtual Mouse, Hand Gesture Recognition, Touchless Interface, Real-Time System.

I.INTRODUCTION

The interaction between humans and computers has traditionally involved physical interaction with computers using devices like keyboards and mice. Even though these devices offer an efficient way of controlling computers, they require physical interaction with computers and do not allow for natural interaction between humans and computers. With the rapid advancement of vision-based technologies, researchers have started exploring alternative interaction techniques between humans and computers that allow touchless interaction with computers using more intuitive control techniques. Hand gesture recognition is one such interaction technique between humans and computers.

Hand gesture recognition has recently attracted considerable attention from researchers due to its potential applications in virtual environments, smart interfaces, assistive technologies, and interactive systems. “Researchers have conducted several studies emphasizing the significance of gesture-based interfaces in improving user experience and developing more natural interaction techniques using gesture-based interfaces in Human-Computer Interaction systems [1], [8]. Researchers have explored different techniques like electromyography

sensors, ultrasonic arrays, and vision-based techniques to recognize hand gestures [2], [5]. Among these techniques, vision-based gesture recognition is more popular because it does not require the user to wear any device and can be implemented using a webcam.

Previous studies in the field of gesture recognition have primarily focused on the detection of human body parts and their movements using depth camera devices. For example, pose estimation systems have shown the ability to recognize the positions of the body and hands of a person through images. This has led to the development of various interaction systems that use hand gestures [10]. Recently, perception systems have been proposed to allow for the efficient detection of hand landmarks in real time. For example, the MediaPipe perception system has enabled the development of accurate hand gesture detection systems using a regular camera [11]. These systems have made it possible to design real-time interaction systems that use hand gestures.

Gesture-based interaction systems have been proposed in the field of human-computer interaction as a method of providing users with efficient interaction systems that mimic the way humans communicate. Previous studies have shown that the use of gesture-based interaction systems improves the level of engagement of users in the interaction process. Moreover, the use of hand gestures in interaction systems has shown the ability to improve the level of accessibility of the interaction systems [12], [13]. The use of hand gestures in interaction systems has thus made the field of gesture-based interaction an important area of study in the development of the next generation of interaction systems.

This research is inspired by the achievements of hand gesture-based interaction systems. The proposed study suggests an interface to manage the work of computers using hand movements, which are performed via a web camera. The computer vision will be used to identify the user's hand and the movement of the fingers to operate the computer. The system will therefore offer the users an effective way of managing the workings of the computer using their hands. The suggested system will therefore give users an effective way of communicating with the computer.

II. RELATED WORK

Hand gesture recognition has emerged as a significant area of research for modern interactive systems, especially within the context of the broader field of human-computer interaction. Various researchers have proposed different approaches to facilitate the interaction between humans and computer systems through the use of gestures.

Other researchers have suggested that large-scale surveys on the general field of gesture recognition and its uses might be put forward. Mohamed et al. suggested a comprehensive survey on the general field of real time hand gesture recognition, with the involvement of computer vision and sensor-based methods to identify and detect gestures [1].

Researchers have proposed various approaches to detect and recognize gestures using different sensing modalities. Shin et al. proposed a gesture recognition approach using surface electromyography, which uses the electrical activity of the muscles to detect and recognize gestures performed by the hands of the user [2]. Similarly, Lee et al. proposed a gesture recognition system using stretchable electromyography sensors and graph neural networks to detect and recognize gestures performed by the hands of the user [4]. Although the proposed approaches have high accuracy rates, the proposed systems require the use of sensors attached to the body of the user to detect gestures performed by the hands.

Other studies have concentrated on different approaches to gesture recognition systems. Joo et al. proposed a system for recognizing gestures by employing an ultrasonic array of sensors along with computational techniques to recognize hand movements [5]. Moreover, gloves have been designed to collect precise data on finger movements for various purposes. These approaches provide accurate hand movement recognition. However, the complexity of the system is increased, which may cause inconvenience for the user.

Vision-based approaches to gesture recognition systems have been found to provide a more convenient approach to the problem. These approaches make use of camera-based systems rather than employing various devices to recognize hand movements. Previous studies by Shotton et al. have shown the ability to recognize different parts

of the human body as well as various poses through depth images. These approaches have greatly influenced the development of various modern systems of interaction [10]. More recent approaches, such as MediaPipe, have shown the ability to recognize hand movements through the use of standard camera systems.

Gesture-based interfaces have also been extensively researched in the context of human-computer interaction design as well. Studies have demonstrated that gesture-based interfaces are beneficial in increasing user engagement and provide a more natural way of interacting with a system when compared to using other input devices [8]. Furthermore, modern human-computer interaction design principles focus on designing interactive systems that are more user-centered and allow more intuitive communication between humans and computers [12], [13].

Despite the advancements in gesture recognition systems, it is observed that existing systems are hardware-intensive or require complex hardware components to recognize gestures. However, vision-based systems using regular cameras provide a more viable solution for designing gesture-based systems in real-time. Keeping this in view, the proposed system aims to design a virtual mouse system using hand gesture detection using computer vision techniques.

III. METHODOLOGY

The proposed system incorporates a vision-based virtual mouse concept through which the user can interact with the computer using hand gestures recognized by the webcam. This system replaces the physical mouse with specific hand gestures to perform the operations of the mouse. The methodology for the proposed system includes various steps, namely image acquisition, hand detection, landmark extraction, gesture recognition, and finally the execution of the operation.

1.System Overview

The virtual mouse system follows a sequential workflow in which the webcam captures video frames. These video frames are processed to identify hand gestures. The hand gestures are used to send commands to the cursor. The main aim of the system is to provide a touchless experience for the user to operate the computer.

Fig 1 : System Architecture

2. Image Acquisition and Pre-processing

The first stage in the process is the acquisition of video feeds in real time using the computer's webcam. Each image obtained from the video feed is processed separately to recognize the hand gestures. Using OpenCV libraries, video feeds are captured and converted into appropriate color formats.Pre-processing operations are performed on the image to increase efficiency in the detection process and minimize complexity in the process. This includes resizing the image, changing the color format if needed, and preparing the image for the next stage in the process—hand tracking.

3. Hand Detection and Landmark Identification

The second stage would be to detect the hand and identify the landmark points. This can be done with the help of MediaPipe. MediaPipe has a powerful hand tracking feature that can detect up to 21 landmarks on the hand. These landmarks include the tips of the fingers and the joints as well as the wrist. Each of these landmarks has spatial coordinates that can be used to identify the position of the hand. By analyzing these coordinates, the position of the hand can be identified. This landmark detection is done for every frame of the video that is captured through the webcam.The landmark points play a critical role in identifying the position of the fingers as well as the state of the fingers.

4. Gesture Recognition Mechanism

The gesture recognition in the proposed system utilizes rule-based logic instead of machine learning. The system

compares the position of the landmarks to identify the gesture that the system is currently performing. For instance, the system compares the distance between certain landmarks on the fingers to identify if the finger is raised or bent. The system can then identify the gesture that the system is currently performing. This approach makes the system simple and does not require training the model.The system defines various gestures to execute mouse-related functions such as moving the cursor, clicking the mouse, scrolling the mouse, and other system functions. The gesture recognition occurs in real time as the hand moves within the field of view of the camera.

5. Cursor Movement Control

This works by tracking the position of your index finger. The position of the tip is then translated into the screen coordinates of your display. This means that as you move your finger in front of the camera, it will move across the screen in response.To make it even better, as well as prevent any shock, there is usually some form of smoothing added to the movement of the cursor. This is to keep it as smooth as possible.

6. Mouse Action Implementation

Various hand movements correspond to mouse functions, depending on the positioning of the fingers. This system translates finger landmark patterns into functional actions.

Some of the common hand movements include:

- Moving the cursor by sliding the index finger.
- Clicking on the left mouse button by touching the index finger and the thumb.
- Clicking on the right mouse button by combining the fingers.
- Double-clicking by repeating the action.
- Scrolling by moving the fingers up and down.
- Adjusting the volume by specific movements.
- Adjusting the brightness by specific finger postures.
- Moving between opened windows by hand movements.

7. Real-Time System Execution

To ensure that everything is running smoothly, the whole system is running nonstop in real time. Every frame from the webcam is being processed in just milliseconds for detection, gesture recognition, and action execution. By utilizing efficient computer vision libraries like OpenCV and taking advantage of the optimized hand tracking from MediaPipe, the system is capable of running in real time.The system is always keeping watch of hand gestures and is always keeping the cursor and actions running in real time. This makes the virtual mouse behave almost like a regular hardware mouse.

IV. PERFORMANCE ANALYSIS

The efficiency of the proposed virtual mouse system was evaluated to assess its efficiency, responsiveness, and reliability in real-time interaction between humans and computers. The evaluation of the proposed virtual mouse was based on certain parameters. The experiments were conducted using a webcam, as well as a computer system capable of running real-time video processing applications.

1.Gesture Detection Accuracy

Accuracy of gesture detection is the ability of the system to recognize and perform the predefined hand gestures and map them to mouse actions. The proposed system uses hand landmark detection for finger position detection and rule-based logic for gesture detection.Experimental observations show that simple hand gestures like cursor movement and single clicks have high accuracy in terms of detection, as they are based on the detection of finger positions. Complex hand gestures involving finger combinations for actions like scrolling and switching windows are effective when the hand is clearly visible in the camera frame.

However, slight inaccuracies in the system may be observed when the hand is not clearly visible in the frame due to partial occlusion or when the objects in the background are many.

2. System Response Time

System response time is defined as the time taken between the user's gesture and the corresponding action of the mouse cursor on the screen. Since the system is dealing with live video feed, efficient frame processing is required for the system's performance. The efficient frame processing of the system is due to the use of optimized computer vision libraries and the real-time hand tracking capability of the system using the MediaPipe library. In most cases, the cursor movement is efficient and provides a similar experience for the user as using a normal mouse.

3. Real-Time Processing Capability

The capability of the system to perform real-time processing is significant in the context of a gesture-controlled system. The proposed system is capable of performing several operations concurrently. These operations include the acquisition of frames from the webcam, hand landmark detection, finger position analysis, and mouse actions. However, the system is capable of maintaining stability in continuous operation. The hand tracking module of the system is efficient in detecting hand landmarks from the webcam feed. The gesture recognition module of the system can perform finger movement analysis without incurring significant computational costs.

4. Usability and User Experience

The usability of the proposed system was tested based on the ease with which the system can be used for performing mouse actions using hand gestures. The proposed system can be used for moving the cursor, clicking the mouse, and scrolling the page without the need for any physical contact with the mouse.The proposed system can be used for providing a natural and intuitive interface for the user. The system can be used for performing actions using finger movements, thereby reducing the need for using hardware devices for input.

5. Environmental Factors

Environmental factors, including lighting and background, might affect the efficiency of vision-based gesture recognition systems. The experimental analysis revealed that the proposed system works efficiently in well-lit conditions and when the background is simple. In cases of low illumination or complex backgrounds, there might be occasional detection errors in hand detection. The limitations of hand detection highlight the importance of a well-controlled environment for optimal efficiency of the proposed system.

6. Overall System Efficiency

Overall, it can be concluded that the proposed virtual mouse system is efficient in performing its functions in real-time. The proposed virtual mouse system is efficient in performing its functions in real-time. The accurate detection of hand poses, combined with rule-based gesture recognition, enables the proposed virtual mouse system to efficiently perform mouse functions.

V. RESULTS AND DISCUSSION

The project was successfully implemented to facilitate touchless interaction with the computer. The project utilizes a webcam to access video feed in real time. The OpenCV library was employed for processing the video feed, while the MediaPipe library was used to access the hand landmark detection results. The hand tracking feature detects 21 critical landmarks on the back of the hand. Analyzing these landmarks enables the identification of various gestures that can be mapped to different mouse functionalities such as moving the mouse pointer, clicking the mouse button, scrolling the mouse wheel, among others.

The mouse functionalities were successfully implemented as the project was able to identify the various hand gestures and respond accordingly. The mouse pointer movement was controlled by the movement of the index finger. When the index finger moved within the camera's view, the mouse pointer moved accordingly. Clicking the mouse button was facilitated through various finger combinations. For instance, the combination of the index finger and the thumb was responsible for the left-click functionality. The combination of the other fingers was responsible for the right-click functionality. Scrolling the mouse wheel was facilitated through the vertical movement of the fingers. The project was responsive to the various hand gestures, allowing the mouse functionalities to be executed within the shortest time possible.

VII. CONCLUSION

In this project, we have designed a virtual mouse that allows you to operate your computer using gestures. A webcam observes the movements of your hands and translates them into mouse movements. We have used computer vision in this project. It has identified the gestures of your hands and has translated them into mouse movements. It has the functionality of moving the mouse pointer, left-clicking, right-clicking, double-clicking, scrolling, brightness control, volume control.

It has been implemented using OpenCV and Mediapipe. It has allowed us to track the movements of the fingers of

the hands and has identified the gestures. Python has helped us integrate these technologies and has allowed us to develop this project.

The project has been successful in identifying the movements of the mouse and has performed most of the mouse movements accurately. It has allowed the mouse pointer to move in the directions of the fingers of the hands. It has identified the gestures of the hands correctly in most of the cases.

However, the project has some limitations. It has performed poorly in low light conditions. It has difficulty in identifying the gestures of the hands when the vision of the webcam is poor. It has performed poorly when the hands have made fast movements". Thus, this project has demonstrated the potential of computer vision in the field of Human-Computer Interaction. It has allowed us to operate the computer using gestures. It has the potential of becoming a substitute for the mouse in the future.

REFERENCES

[1] Aws Saood Mohamed, Nidaa Flaih Hassan, Abeer Salim Jamil, "RealTime Hand Gesture Recognition: A Comprehensive Review of Techniques, Applications, and Challenges," Cybernetics and Information Technologies, vol. 24, no. 3, 2024. URL: sciendo.com

[2] Shin, J., Miah, A.S.M., Konnai, S. et al. "Hand gesture recognition using sEMG signals with a multi-stream time-varying feature enhancement approach," Scientific Reports,vol.14,no.1,2024.URL:https:// www.nature.com/articles/s41598-024-72996-7

[3] Jungpil Shin, Abu Saleh Musa Miah, Md. Humaun Kabir, Md. Abdur Rahim, Abdullah Al Shiam, "A Methodological and Structural Review of Hand Gesture Recognition Across Diverse Data Modalities," arXiv preprint arXiv:2408.05436, 2024. URL: https://arxiv.org/abs/2408.05436

[4] Lee, H., Lee, S., Kim, J. et al. Stretchable array electromyography sensor with graph neural network for static and dynamic gestures recognition system. npj Flex Electron 7, 20 (2023). URL: nature.com

[5] Joo J, Koh J, Lee H. Hand Gesture Recognition Using Ultrasonic Array with Machine Learning. Sensors.2024;24(20):6763. https://doi.org/10.3390/s24206763. URL: mdpi.com [6] R. Dechter, 'Learning While Searching in Constraint-SatisfactionProblems.', in AAAI, Jan. 1986, pp. 178–185.

[7] Palmerston JB, Zhou Y, Chan RHM. Comparing biological and artificial vision systems: Network measures of functional connectivity. Neurosci Lett. 2020 Nov 20;739:135407.

[8] R. P. Sharma and G. K. Verma, 'Human Computer Interaction using Hand Gesture', Procedia Computer Science, vol. 54, pp. 721–727, 2015, doi: 10.1016/j.procs.2015.06.085.

[9] Jeremy Chan, Eduardo Veas, Jörg Simon, Designing a Sensor Glove Using Deep Learning, IUI '21: Proceedings of the 26th International Conference on Intelligent User Interfaces, pp 150 – 160. https://doi.org/10.1145/3397481.3450665

[10] J. Shotton et al., 'Real-Time Human Pose Recognition in Parts from a Single Depth Image', Jun.2011.[Online].Available:https://www.microsoft.com /en-us/research/publication/real-time-humanpose-recognition-in-parts-from-a-single-depth-image/

[11] C. Lugaresi et al., 'MediaPipe: A Framework for Building Perception Pipelines', arXiv:1906.08172 [cs], Jun. 2019, Accessed: Nov. 26, 2021. [Online]. Available: http://arxiv.org/abs/1906.08172

[12] A. Skarlatidou and C. I. Otero, 'Design approaches and human– computer interaction methods to support user involvement in citizen science'. Book Chapter, UCL Press., 2021

[13] J. T. George & M. J. George, 'Human-Computer Interaction Tools and Methodologies' Human-Computer Interaction in Game Development with Python: Design and Develop a Game Interface Using HCI Technologies and Techniques (pp. 1–34). Springer, 2022