International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 Impact Factor: 8.176 ISSN: 2582-3930

Gesture Controlled Virtual Mouse

Dr. N Kumaran, Bathula Shanmukha Tarun, Adusumalli Hari Krishna

Assistant Professor, Dept. Of CSE, SCSVMV (Deemed to be University)

Student, Dept. Of CSE, SCSVMV (Deemed to be University)

Student, Dept. Of CSE, SCSVMV (Deemed to be University)

Abstract - Recent advancements in hand tracking and gesture recognition have created both possibilities and difficulties. The field of computer vision is becoming more popular, and technology that is readily available and capable of supporting new advances is being developed and improved quickly. This article discusses and puts into practice some of these options while summarizing difficulties and promising directions for the future of virtual reality and human user interaction. Given the popularity of COVID19, the aim behind this dissertation is to decrease human interactions and reliance on gadgets to operate computers. These findings will encourage more study and, in the long term, support the usage of virtual environments. There are no such limitations in the suggested period, which may instead rely on gesture recognition. During this expedition, activities like A variety of hand motions may be used to click and drag things. Only a camera will be needed for the proposed project's input method. Python and OpenCV are the two software tools needed to implement the suggested machine. The output from the digital camera may be seen on the machine's screen so that the user can calibrate it in addition. NumPy, math, wx, and mouse are the Python requirements that will be utilized to create this machine.

Key Words: Gesture control, Media pipe, OpenCV.

INTRODUCTION

The users who lack a physical mouse can still operate their computer by using a virtual mouse. It may be regarded as hardware because it makes use of a typical webcam. A virtual mouse can be used in conjunction with input devices such as a real mouse or a computer keyboard. The virtual mouse driven by a camera employs several image processing techniques. The user's hand motions are translated into mouse clicks. A web camera's default setting is for continuous photo capture. Recently, webcam-equipped computers have started using facial recognition security software. This may be accomplished by employing the webcam's vision-based CC feature, which does away with the need for computer mice and keyboards. Various other HCI applications, such as motion controllers, databases, or sign language, may also gain from employing a camera. Two recent developments in HCI gaming technology are the Microsoft Kinect and the Nintendo Wii. The enjoyment and interactivity of playing video games has increased as a result of this new gaming technology. The Nintendo Wii is a prime illustration of how motion controllers change gaming, having sold more than 50 million copies in its first year. Hand gestures are incredibly intuitive and efficient for one-on-one computer interaction, and they provide a Natural User Interface (NUI). Cursor control with hand gestures has been thoroughly researched. New tools and procedures have been created. Hand gesture detection is crucial because it distinguishes sign languages and HCI.

LITERATURE SURVEY

In the current virtual mouse control system, we can operate the mouse pointer, left click, right click, drag, and other basic mouse functions using a hand recognition system. In the future, hand recognition won't be utilized. Despite the fact that there are numerous systems for recognizing hands, the system they employed is static hand recognition, which merely recognizes the shapes that hands make and defines an action for each shape, which is constrained to a small number of defined actions and leads to considerable confusion. There are more and more alternatives to utilizing a mouse as technology develops. Gesture Controlled Virtual Mouse uses voice commands and hand gestures to make using a computer with a human being simple. The amount of direct touch with the computer is minimal. Static and dynamic hand motions, together with a voice assistant, may virtually handle all i/o activities. This project uses cutting-edge ComputerVision and Machine Learning techniques to identify hand movements and vocal instructions, and it operates without the need for any extra hardware. It makes use of models like CNN that Media Pipe, which runs on top of pybind11, has created. We'll create a

gesture-controlled virtual mouse using ESP32-CAM & OpenCV. The wireless mouse tracking and clicking functions may be managed using the ESP32 Camera Module and a Python application.

One has to have a solid understanding of Python, image processing, embedded systems, and the Internet of Things in order to get started. We will first comprehend how to manage mouse tracking and clicking as well as all of the prerequisites needed to start a Python application. We will first use a webcam or internal camera from a laptop to test the entire Python script. The Python code will be executed in the second section using an ESP32-CAM Module. Input will therefore be provided via the ESP32-CAM rather than a PC camera or any other external camera.

The main goal of the proposed AI virtual mouse system is to create a replacement for the conventional mouse system that can perform and control mouse functions. This can be done with the aid of a web camera that records hand gestures and hand tips and then processes these frames to perform the specific mouse function, such as the left click, right click, and scrolling function. Gesture Controlled Virtual Mouse uses voice commands and hand gestures to make user-computer interaction simple. The amount of direct touch with the computer is minimal. Static and dynamic hand motions, together with a voice assistant, may virtually handle all i/o activities. Modern machine learning and computer vision techniques are used in this project to identify hand gestures and vocal instructions, and they function without the need for any extra hardware. It makes use of models like CNN that Media Pipe, which is built on top of pybind11, has provided. It comprises of two modules: one that uses Media Pipe Hand detection to operate directly on hands, and the other that uses gloves of any uniform color. It currently runs on the Windows operating system.

Table -1: Comparison table

S. No	Name of the journal	Methods Used	Accuracy
1.	Virtual Mouse using Hand Gestures	Human Computer Interaction, Python, Color Detection, Web camera, Hand Gestures.	82%
2.	Virtually controlling computer using hand-gestures and voice	CNN implemented by Media Pipe running on top of pybind11, OpenCV python.	89%
3.	Gesture Controlled Virtual Mouse with ESP32-CAM & OpenCV	ESP32-CAM & OpenCV	83%

PROPOSED SYSTEM

To create a virtual mouse that can only be used with the fingers and can recognize hand motions. In accordance with which particular combination of fingers is detected, we employed numerous combinations of fingers to carry out various mouse actions. The usage of a virtual mouse can be seen when there is a need to save space or when moving about. Users of the proposed system are not forced to utilize any devices or sensors, and they are not compelled to paint their fingertips with a certain color. It enables user interaction with a computer without the need for a hardware mouse controller. Additionally, it is inexpensive and simple to use. The OpenCV library is used for computer vision tasks, while the Media Pipe framework is used for tracking and detecting hand movements. The application uses machine learning principles to monitor and distinguish between hand motions and hand tips. A framework called Media Pipe is a Google open-source framework that is applied in a machine learning pipeline. Since the Media Pipe framework was created utilizing time series data, it may be used for cross-platform programming. The Media Pipe architecture supports several audio and video formats since it is multimodal. The Media Pipe framework is used by the developer to create and analyze systems using graphs as well as to create systems for application-related purposes. The actions in the Media Pipe-using system are performed in the pipeline configuration. Scalability on desktops and mobile devices is made possible by the pipeline's flexibility to execute on several platforms.

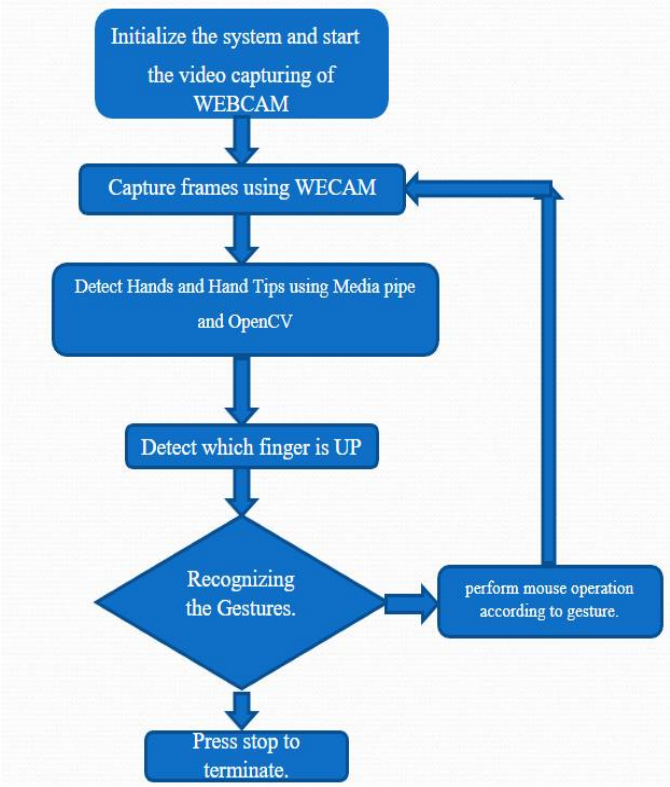


Fig -1: Figure

Algorithm

Step 1: Start the program.
Step 2: Open the file and using that file location go to cmd.

Step 3: Using some libraries installed for the code run the program.

Step 4: Initialize the system and start the video capturing of WEBCAM.

Step 5: Capture frames using WECAM.

Step 6: Detect Hands and Hand Tips using Media pipe and OpenCV.

Step 7: Detect which finger is UP.

Step 8: Recognizing the Gesture.

Step 9: Perform mouse operations according to gesture.

Step 10: Stop the program.

METHODOLOGY

The transformational method is used by the gesture-based virtual mouse system to translate the fingertip's coordinates from the camera screen to the full-screen computer window for operating the mouse. A rectangular box is generated in relation to the computer window in the camera zone where we move the mouse pointer around the window when the hands are identified and we determine which finger is capable of completing the specified mouse operation.

Modules:

- OpenCV
- Media Pipe
- PyAutoGUI
- Math

A. OpenCV

A computer vision package called OpenCV contains techniques for processing images that detect objects. Real-time computer vision applications may be made using the Python computer vision package known as OpenCV. The OpenCV library is used to analyze data from photos and videos, including face and object detection. A free and open-source software library for computer vision and machine learning is called OpenCV. A standard infrastructure for computer vision applications was created with OpenCV in order to speed up the incorporation of artificial intelligence into products. OpenCV makes it simple for companies to use and alter the code because it is a product with an Apache 2 license.

B. Media Pipe

A framework called Media Pipe is a Google open-source framework that is applied in a machine learning pipeline. Since the Media Pipe framework was created utilizing time series data, it may be used for cross-platform programming. The Media Pipe architecture supports several audio and video formats since it is multimodal. The Media Pipe framework is used by the developer to create and analyze systems using graphs as well as to create systems for application-related purposes. The pipeline configuration is where the actions in the Media Pipe-using system are carried out. Scalability on desktops and mobile devices is made possible by the pipeline's flexibility to execute on several platforms. The three essential components that make up the Media Pipe framework are performance evaluation, a system for accessing sensor data, and a group of reusable pieces known as calculators. A pipeline is a graph made of units called calculators that are connected to one another by streams via which data packets pass. In order to create their own application, developers can add, remove, or redefine custom calculators anywhere in the graph.

C. PyAutoGUI

In essence, PyAutoGUI is a Python software that runs on Windows, MacOS X, and Linux and allows users to imitate keyboard button pushes as well as mouse cursor movements and clicks. A Python package for cross-platform GUI automation for people is called PyAutoGUI. A Python automation module called PyAutoGUI may be used to click, drag, scroll, move, etc. It may be used to click precisely where you want. used to automate the control of the keyboard and mouse. There are several techniques to programmatically control the mouse and keyboard in each of the three main operating systems (Windows,

macOS, and Linux). This frequently involves complex, enigmatic, and highly technical elements. PyAutoGUI's role is to conceal all of this complexity behind a straightforward API.

D. Math

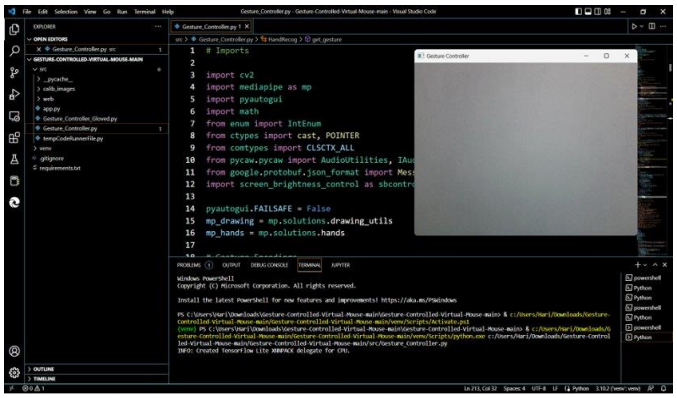
A crucial component created to deal with mathematical operations is the Python math module. It has always been there and is supplied with the default Python version. The majority of the mathematical functions in the math module are only thin wrappers for the C platform's equivalents. The math module is effective and adheres to the C standard since its underlying functions are built in CPython. You have the option to carry out frequent and practical mathematical operations inside your program thanks to the Python math module. There are several preset constants available in the Python math module. Access to these constants has a number of benefits. One benefit is that you may save a lot of time by not having to manually hardcode them into your program. Additionally, they provide uniformity across your whole code.

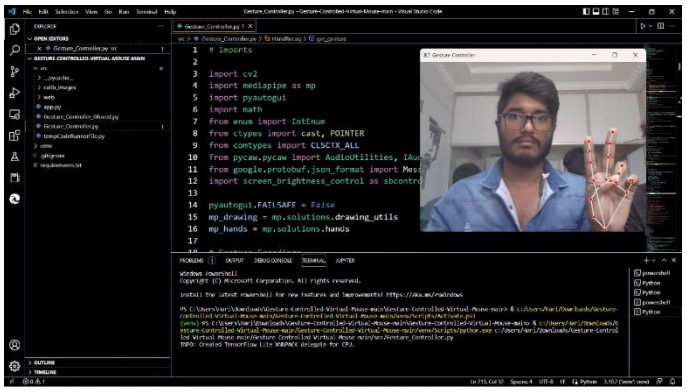
PROJECT DESCRIPTION

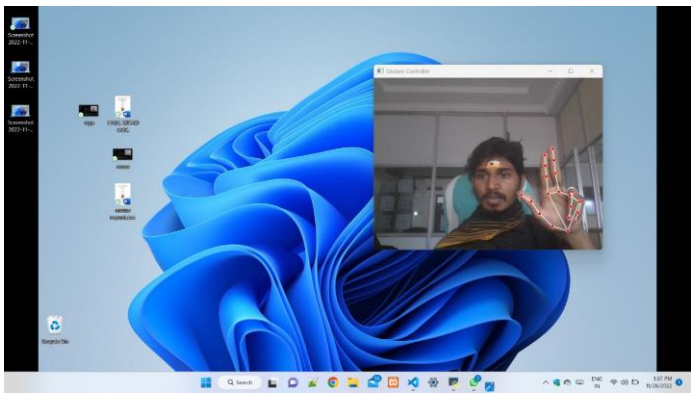
Human-computer contact is getting more and more convenient in daily life as the usage of computers has been ingrained in our lives. While most people take these places for granted, persons with impairments have a lot of trouble using these things correctly. This paper introduces a gesture-based virtual mouse system that employs hand movements and hand tip recognition to simulate mouse operations in a computer. The major objective of the recommended system is to replace the traditional mouse with a web camera or a computer's built-in camera to perform computer mouse pointer and scroll functions. Real-time detection and identification of a hand or palm is accomplished using a single-shot detector model. The Media Pipe uses the single-shot detector concept. Because it is simpler to learn palms, the hand detection module initially trains a model for palm detection. Furthermore, for tiny objects like hands or fists, the Non maximum suppression performs noticeably better. Finding joint or knuckle coordinates in the hand area is a model for a hand landmark. The camera frames from a laptop or PC serve as the foundation for the planned Gesture virtual mouse system. The video capture object is constructed using the Python computer vision package OpenCV, and the web camera begins recording footage. The virtual system receives frames from the web camera and processes them. The transformational method is used by the gesture-based virtual mouse system to translate the fingertip's coordinates from the camera screen to the full-screen computer window for operating the mouse. A rectangular box is generated in relation to the computer window in the camera zone where we move the mouse pointer around the window when the hands are identified and we determine which finger is capable of completing the specified mouse operation.

RESULT

As computer use has been engrained in our everyday lives, human-computer interaction is becoming more and more convenient. While most people take these areas for granted, people with disabilities frequently struggle to use them properly. In order to imitate mouse activities on a computer, this study offers a gesture-based virtual mouse system that makes use of hand motions and hand tip detection. The main goal of the suggested system is to swap out the conventional mouse for a web camera or a built-in camera on a computer to perform mouse pointer and scroll tasks.







CONCLUSIONS

The primary goal of the virtual mouse system is to replace the usage of a physical mouse with the use of hand gestures to control mouse cursor functions. The suggested system may be implemented utilizing a webcam or an integrated camera that recognizes hand movements and hand tips and analyses these frames to carry out certain mouse actions. The model has certain drawbacks, including a little loss of precision in right-click mouse functionality and some challenges with dragging and clicking to pick text. In order to get over these restrictions, we will now concentrate on making the fingertip detection algorithm more precise.

REFERENCES

1. Rao, A.K., Gordon, A.M., 2001. Contribution of tactile information to accuracy in pointing movements. Exp. Brain Res. 138, 438–445. https://doi.org/10.1007/s002210100717.
2. Masurovsky, A., Chojecki, P., Runde, D., Lafci, M., Przewozny, D., Gaebler, M.,2020. Controller-Free Hand Tracking for Grab-and Place Tasks in Immersive Virtual Reality: Design Elements and Their Empirical Study. Multimodal Technol.Interact.4,91. https://doi.org/10.3390/mti4040091.
3. Lira, M., Egito, J.H., Dall'Agnol, P.A., Amodio, D.M., Gonçalves, Ó.F., Boggio, P.S., 2017. The influence of skin colour on the experience of ownership in the rubber hand illusion. Sci. Rep. 7, 15745. https://doi.org/10.1038/s41598-017- 16137-3.
4. Inside Facebook Reality Labs: Wrist-based interaction for the next computing platform [WWW Document], 2021 Facebook Technol. URL https://tech.fb.com/inside-facebook-realitylabs-wrist-basedinter action-for-the-next computing-platform/ (accessed 3.18.21)
5. Danckert, J., Goodale, M.A., 2001. Superior performance for visually guided pointing in the lower visual field. Exp. Brain Res. 137, 303–308. https://doi.org/10.1007/s002210000653.
6. Carlton, B., 2021. HaptX Launches True-Contact Haptic Gloves For VR And Robotics. VRScout. URL https://vrscout.com/news/haptx-truecontact-haptic-gloves-vr/ (accessed 3.10.21).
7. Brenton, H., Gillies, M., Ballin, D., Chatting, D., 2005. D.: The uncanny valley: does it exist, in: In: 19th British HCI Group Annual Conference: Workshop on Human-Animated Character Interaction.
8. Buckingham, G., Michela kakis, E.E., Cole, J., 2016. Perceiving and acting upon weight illusions in the absence of somatosensory information. J. Neurophysiol. 115, 1946–1953. https://doi.org/10.1152/jn.00587.2015.
9. J. Katona, "A review of human–computer interaction and virtual reality research fields in cognitive Info Communications," Applied Sciences, vol. 11, no. 6, p. 2646, 2021.View at: Publisher Site | Google Scholar.
10. D. L. Quam, "Gesture recognition with a DataGlove," IEEE Conference on Aerospace and Electronics, vol. 2, pp. 755–760, 1990. View at: Publisher Site | Google Scholar
11. D.-H. Liou, D. Lee, and C.-C. Hsieh, "A real time hand gesture recognition system using motion history image," in Proceedings of the 2010 2nd International Conference on Signal Processing Systems, IEEE, Dalian, China, July 2010. View at: Publisher Site | Google Scholar
12. S. U. Dudhane, "Cursor control system using hand gesture recognition," IJARCCE, vol. 2, no. 5, 2013. View at: Google Scholar.
13. K. P. Vinay, "Cursor control using hand gestures," International Journal of Critical Accounting, vol. 0975–8887, 2016. View at: Google Scholar.
14. J. Katona, "A review of human–computer interaction and virtual reality research fields in cognitive Info Communications," *Applied Sciences*, vol. 11, no. 6, p. 2646, 2021.
15. J. Jaya and K. Thanushkodi, "Implementation of classification system for medical images," European Journal of Scientific Research, vol. 53, no. 4, pp. 561–569, 2011. View at: Google Scholar.

BIOGRAPHIES

Dr. N. Kumaran received his Ph.D. from the National Institute of Technology, Tiruchirappalli., M. Tech (Information Technology) from Sathyabama University, Chennai, and B.E (Computer Science and Engineering) from Coimbatore Institute of Technology, Coimbatore-14, Tamil Nadu, India, in the year 2007 and 1998 respectively. He is currently working as an Assistant Professor in the Department of Computer Science and Engineering, Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya (Deemed to be University) Kanchipuram. He has served as a reviewer committee for an international journal. He has been an active member of professional bodies, such as IAENG and the Judge of Toycathon 2021 (Ministry of Education and Innovative Cell). He also received the best teacher award in 2016 at his university. His areas of interest include Video Analysis, Machine Learning, the Internet of Things, and Computer Networks. He has got around 20 years of teaching experience in various Institutions.

Bathula Shanmukha Tarun is in his final year of Bachelor of Engineering in the department of Computer Science and Engineering (CSE) at Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya Deemed to be University, Enathur, Kanchipuram, Tamil Nadu, India. His areas of interest include Full Stack Web Development, Machine Learning, and Data Science.

Adusumalli Hari Krishna is in his final year of Bachelor of Engineering in the department of Computer Science and Engineering (CSE) at Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya Deemed to be University, Enathur, Kanchipuram, Tamil Nadu, India. His areas of interest include Full Stack Web Development, Blockchain Technology.