Gesture-Controlled Virtual Mouse using Media Pipe

Samatha Konda[1], Dr. P. IlaChandana Kumari[2], Dr. P.Madhuri[3], Chintolla Surekha[4], *Bhargavi Thatikonda[5]

[1]Assistant Professor, MuffakamJah Engineering College, Panjagutta, Hyderabad
[2,3]Associate Professor, Hyderabad Institute of Technology and Management, Hyderabad
[4]Assistant Professor, Hyderabad Institute of Technology and Management, Hyderabad
[5]Assistant Professor, St. Martin's Engineering College, Secunderabad, Telangana-500100
Email: tbhargaviit@smec.ac.in

ABSTRACT

A Human-Computer Interaction software that enables the user to interact with the computer just by the means of Hand-Gestures. The proposed system uses Machine Learning based Packages like Media Pipe in which a model named "Single Shot Detector" in the package has trained to detect the desired objects from the camera stream provided by OpenCV, It captures the hand gestures and the program triggers specific mouse function accordingly. This offers the user to carry multiple Mouse functions and also to access a few System level controls, It allows us for "Left, Right, and Double Click functions", "Scrolling", "Cursor Navigation", "Drag and Drop", "Selection", "Volume Up", "Volume Down", "Brightness Up", and "Brightness Down". It can also be used in times of pandemics, and few conditional use cases where we don't have an option to hold a mouse such as in presentations while wearing Virtual Reality Headsets, etc. In this paper, we had proposed a revolutionary model that let the users to control their machines without any physical (external).

Keywords: Machine Learning, Open CV, HCI.

I. INTRODUCTION

Gestures are used to communicate nonverbally and to deliver a certain message. This message can be sent through a person's body, hands, or face movements. When interacting with others, gestures can be used to express information from easy to highly difficult hand motions. For illustrative exa mple, we can employ straightforward gestures or motions that are expressed in sign languages and are included into their syntax to point to anything (a people or object), or employ range of many ot her simple gestures or motions. As a result, employing hand gestures as a tool, humans can engage with one another more efficiently with the aid of computers.

The movement of a visual object is one mouse function that has been replaced by hand movemen ts. The work is designed to be cheap, and it captures hand gesture via a cam, one of many cheap in put devices. Preset command-based movements are modeled to touch materials. There are numero us current systems. One can move around the monitor using a standard mouse (hardware tool). Th e monitor screen cannot be accessed with hand gestures. Another is the gesture system, which reco gnizes gestures using colored tapes. Additionally, the functions are static and simple in nature. Usi ng current technique, we could operate the mouse and do some basic tasks on a computer or lapto p with a web camera and microphone without the need of any other computer hardware. Other pro cedures can be done with a voice assistant.

II. LITERATURE REVIEW

Hardware-based system, by Quam [1] tells that with a human hand operating the Data Glove, an e xperiment was carried out to look at gesture recognition. In three classes, a total of 22 gestures we

re examined. Only finger flexure movements were used in the first class. There were movements i n the second class that called for both finger flexion and hand orientation. The third class of gestur es need finger motion in addition to flexure and orientation. The flex sensors which are present in t he data glove make it easy to recognize up to 15 different movements.

Undoubtedly, the precise sensors required rely on the motions themselves as well as the kind of ge stures that need to be identified. Although it appears that class 3 gestures could be recognized, add itional research will be required to create recognition techniques.

Dung-Hua Liou and Chen-Chiung Hsieh proposed A Real Time Hand Gesture Recognition Syste m Uses Motion History Image [2]. In this study, a face based skin color model and a motion histor y image based hand movement detection technique were developed. Hands travelling up, down, le ft, right are the dynamic hand gestures as well as two static hand gestures the fist and the wave ha nd are suggested in this work. These hand gestures are simple and effortless. Using Harr-like featu res, the four-directional dynamic hand movements were identified. Static hand movements were e xtracted using the face-based adaptive skin colour model, and these motions were subsequently re cognised by examining a face-based ROI. To test our suggested system, five people were invited. Experimental results revealed an average accuracy of 94.1% and proved the viability of the sugges ted approach.

Thumma Dhyanchand, Vantukala VishnuTeja Reddy came up with a Virtual Mouse Control Usin g Colored Finger Tips and Hand Gesture Recognition the system, [3] that makes the control of a c ursor without any direct physical contact without any sensor. This activity involves identifying col ourful fingertips and tracking them. For the same effect, different hand gestures might be used in p lace of colored caps.

The mouse can be used to scroll, click once on the left side, click twice on the left side, and do oth er functions. For various processes, different arrangements of the colored caps are employed. Dep ending on the person being utilized and the lighting environment, the application can change the r ange of skin tones. After examining the program output at various hand motions, an approximatio n of the area ratio that the hand is not using in the convex hull is made. As the brightness ranges fr om 500 to 600 lux the color Red has a detection accuracy around 90% which is similar in case of Green and Blue which is typical of offices and well-lit classrooms. This problem is solved by adap ting a hand gesture recognition technology that detects the contours of the hand.

An Introduction to Hidden Markov Models by *L. R. Rabiner B. H. Juang* [4] tells that a key tool f or the real-time, dynamic gesture identification process is the Hidden Markov Model. The HMM a pproach is practical and designed to function in static settings. The strategy involves using the HM M's LRB topology in co-occurrence with Baum Welch Algorithm and Forward and Viterbi Algori thms for training and testing respectively, produces the best recognition of patterns. Although the s ystem in this study looks to be simpler to use than more recent systems or command-based system s, it is less effective at spotting and recognizing patterns. An Arduino Uno, ultrasonic sensors, and a laptop are used in this study's hand gesture laptop to carry out tasks including managing media p laying and volume. Serial connections are made using Python, Arduino, and ultrasonic sensors. Fo r more engaging and interactive learning, immersive gaming, and interacting with virtual things on screen, this kind of technology can be used in the classroom.

Akshaya U Kulkarni, Amit M Potdar proposed a system that is RADAR based Object Detector usin g Ultrasonic Sensor [5]. The project entailed developing an ultrasonic sensor based, RADAR based used for object detection, was provided in this publication. Instead of employing genuine RADAR, which is expensive and difficult to handle, it provides a solution for simple object detection using ul trasonic technology that functions like RADAR. The work of other authors focuses primarily on eith er of these subjects. IoT hardware and connection software were part of the endeavor. The Raspberr

y Pi 3 computer and Arduino Uno board processed data. In order to identify objects, the boards were equipped with an ultrasonic sensor and servo motor. The SIM808 module was then used to send eac h object's distance, angle, and timestamp to the chosen number via SMS/message. Sample test cases were included in the results to verify the object detection's detection range. The study provides a sim ple approach for object detection since, as stated in the introductory part, ultrasonic detection has va rious benefits over RADAR.

D.Ghosh, P.K.Bora, and M.K.Bhuyan Co- articulation Detection in Hand Gestures [6] suggests th at one of the biggest problems with dynamic gesture recognition is co-articulation. For the class of gestures taken into consideration here, there haven't been many documented vision-based methods for assessing co-articulation. The majority of the algorithms that have been suggested up to this po int have been successful only for a small number of gesture vocabularies and cannot be applied to all types of gestures when used in various settings. Another significant issue with dynamic gesture identification is the self-articulation of gesture in the sequence of gestures. When employing the provided method for recognizing co- articulation, the connected gesture sequences in the gesture v ocabulary that are used in light of some particular applications, such as robotic control etc.

Deep Learning-Based Real-Time Artificial Intelligence Virtual Mouse by S. Shriram, B. Nagaraj [7] states that by using a built-in camera or a webcam that recognizes hand movements and finger-t ips and frames are detected to carry out certain mouse actions. The model's outcomes demonstrate the suggested AI virtual mouse system's outstanding performance which is more accurate than the current models, and also overcomes the bulk of the latter's shortcomings. Since the recommended system is more accurate this model can be easily put in practice. With the use of this system the us age of actual mouse can be avoided which reduces the spread of Corona Virus.

On-device Real-time Hand Tracking using Media Pipe by G. R. Fan Zhang, V. Bazarevsky [8] In this post, they recommended Media Pipe Hands, a complete hand tracking system that operates in real- time on different platforms. The pipeline can be easily installed on standard devices and pred icts 2.5D landmarks without the need for specialized hardware. We open sourced the pipeline to e ncourage academics and engineers to create cutting-edge gesture control and AR/VR applications utilizing it.

Real-time virtual mouse system using RGB-D images and fingertip detection by S.-H. Kim, N.-H. Ho, D.-S. Tran, H.-J. Yang, and G. S. Lee [9] this study introduced a novel virtual-mouse techniqu e based on fingertip detection and RGB-D pictures. Just using fingertips in front of webcam user c an perform certain actions. The method showed off not only extreme accurate gesture estimates us eful applications. The proposed approach gets around the drawbacks of the majority of existing vir tual-mouse systems. It has several benefits, including accurate fingertip tracking at a greater distan ce and with complicated backgrounds. It also works well in shifting light conditions. The results o f the experiments showed this method is a better one for real-time hand gesture interfaces.

Hand gesture recognition for human computer interaction by A. Subramanian, A. Haria, J. S. Nay ak N. Asokkumar, and S. Poddar [10] we were able to create a robust gesture recognition system t hat was affordable and simple to use without the usage of any markers. With the help of our gestur e detection technology, we aimed to offer gestures for almost all HCI-related operations, such as system functioning, application activation, and opening a number of well-known websites. Inc reasing accuracy and adding additional movements to incorporate more features are future objecti ves. In addition, we intend to incorporate our tracking system into a variety of hardware, including digital TV and mobile devices, and broaden the possibilities for our domain. Additionally, we w ish to make this mechanism accessible to a wide range of users, including those with impairments.

III. METHODOLOGY

For hand tracking and gesture recognition, the Media Pipe framework is employed, for computer v ision we use the Open CV library. The technique tracks, recognizes hand gestures and finger tips u sing ML concepts.

1. *Open CV:* For images, Open CV offers object detection techniques. Using the Open CV module, the best applications of computer vision can be developed. This library is utilized for face and object identification, picture and video data processing, and video analysis.
2. *Media Pipe*: A machine learning pipeline employs a Google open-source framework known as Media Pipe. The Media Pipe framework may be employed for cross-platform computing as it was developed leveraging time series data. The Media Pipe architecture supports multiple audio and video formats since it is multimodal. The developer uses Media Pipe framework to build systems for application- related objectives as well as to design and analyze systems that employ graphs.

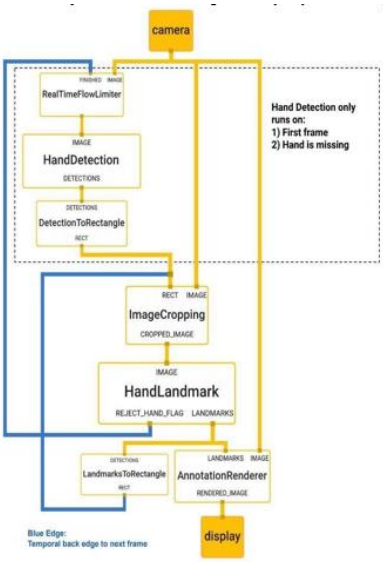
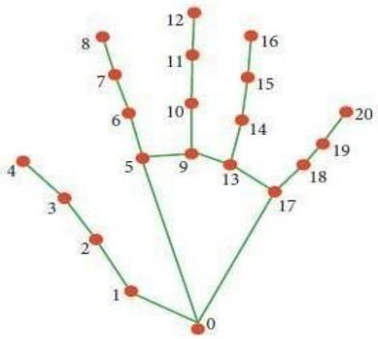


Fig 1: MediaPipe Hand recognition

3. The AI-Powered Virtual Mouse System's Camera: This AI virtual mouse system relies on the im ages from a laptop or PC. The video capture object is generated by using Python computer vision t oolkit Open CV and web camera used for recording. The web cam provides frames to virtual AI s ystem that processes them.

4. Video recording and analyzing: The webcam will be used by the AI virtual mouse system to rec ord each frame till the program is finished. The images are captured and converted to RGB to allo w for frame-by-frame identification of the hands.

5. Virtual Screen Matching: This is used to move the hand coordinates between the web cam and t he computer's window to execute certain mouse functions, the AI virtual mouse technique enables a transformational method. After the recognition of the finger tips and hands we are notified of wh ich fingers are capable of performing a cursor movement, a rectangle box may be generated on the computer screen showing the reference of the web cam. From there, we can see our cursor movem ents around the window.

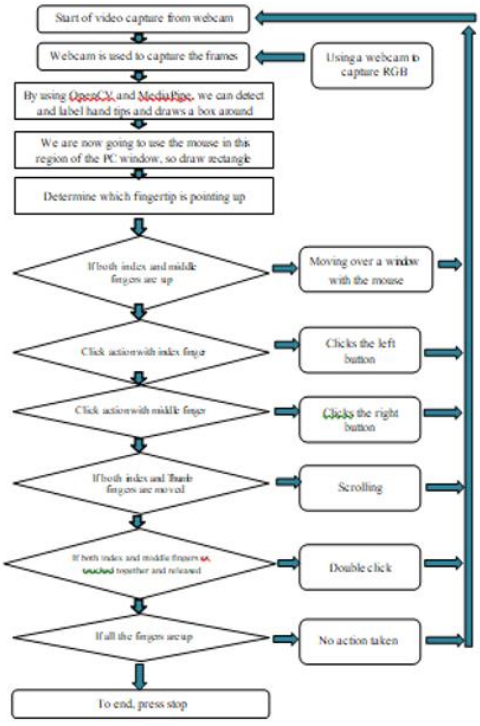


0. WRIST	11. MIDDLE_FINGER_DIP
1. THUMB_CMC	12. MIDDLE_FINGER_TIP
2. THUMB_MCP	13. RING_FINGER_MCP
3. THUMB_IP	14. RING_FINGER_PIP
4. THUMB_TIP	15. RING_FINGER_DIP
5. INDEX_FINGER_MCP	16. RING_FINGER_TIP
6. INDEX_FINGER_PIP	17. PINKY_MCP
7. INDEX_FINGER_DIP	18. PINKY_PIP
8. INDEX_FINGER_TIP	19. PINKY_DIP
9. MIDDLE_FINGER_MCP	20. PINKY_TIP
10. MIDDLE_FINGER_PIP	

Fig 2: co-ordinates in hand

6. Recognizing the finger which is up and performing out the appropriate mouse operation: In ord er to transfer the hand coordinates between the webcam to the computer's window full screen for mouse operation, the AI virtual mouse technique enables a transformational mechanism. After the hands are detected and we are notified of which finger is capable of performing a certain mouse m ovement, a rectangular box is formed in reference to the computer window in the camera region. F rom that, we may move the mouse pointer around the window.

IV. IMPLEMENTATION



V. RESULTS AND EVALUATION

Conclusion

The development of effective human-machine interaction has become significantly impacted by th e hand gesture detection and voice assistant systems. Wide- ranging applications in the technology sector are promised by implementation employing hand gesture recognition. The Media Pipe, a m achine learning framework, has a significant impact on the creation of this application that uses ha nd gesture recognition.

References

[1] D. L. Quam, “Gesture recognition with a DataGlove,” IEEE Conference on Aerospace and Electronics, vol. 2, pp. 755–760, 1990. [2] D.H. Liou, D. Lee, and C.C. Hsieh, “A real time hand gesture recognition sy stem using motion history image,” in Proceedings of the 2010 2nd International Conference on Signal Proc essing Systems, July 2010. [3] V. V. Reddy, T. Dhyanchand, G. V. Krishna and S. Maheshwaram, \"Virtua l Mouse Control Using Colored Finger Tips and Hand Gesture Recognition,\" 2020 IEEE-HYDCON, 2020 , pp. 1- 5, DOI: 10.1109/HYDCON48903.2020.9242677. [4] L. R.Rabiner B. H. Juang “An Introduction to Hidden Markov Models” IEEE ASSP Magazine (Volume: 3, January 1986) [5] Akshaya U Kulkarni, Amit M Potdar “RADAR based Object Detector using Ultrasonic Sensor” 2019 1st International Conference on Advances in Information Technology (ICAIT) 10 February 2020 [6] M.K. Bhuyan, D. Ghosh and P.K. Bor a “Co- articulation Detection in Hand Gestures” TENCON 2005 - 2005 IEEE Region 10 Conference Febru ary 2007. [7] S. Shriram, B. Nagaraj ,“Deep Learning-Based Real-Time AI Virtual Mouse” Volume 2021 A rticle ID 8133076 https://doi.org/10.1155/2021/8133076 [8] V. Bazarevsky and G. R. Fan Zhang. On- Devi ce, MediaPipe for Real-Time Hand Tracking. [9] A. Haria, A. Subramanian, N. Asok kumar, S. Poddar, an d J. S. Nayak “Hand gesture recognition for human compute interaction, ”Procedia Computer Science, vol. 115, pp. 367–374, 2017. [10] D.-S. Tran, N.-H. Ho, H.-J. Yang, S.-H. Kim, and G. S. Lee, “Real-time virtu al mouse system using RGB-D images and fingertip detection,” Multimedia Tools and Applications Multi media Tools and Applications, vol. 80, no. 7, pp. 10473– 10490, 2021.