As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology, 8(01): 26-31, 2022

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0801005
Available online at: http://www.ijmtst.com/vol8issue01.html

Virtual Control Using Hand – Tracking

Dhruv Rustagi | Gaurav Maindola | Harshit Jain | Bhanu Gakhar | Gunjan Chugh

Department of Information Technology, Dr. Akhilesh Das Gupta Institute of Technology and Management, New Delhi, India.

To Cite this Article
Dhruv Rustagi, Gaurav Maindola, Harshit Jain, Bhanu Gakhar and Gunjan Chugh. Virtual Control Using Hand – Tracking. *International Journal for Modern Trends in Science and Technology* 2022, *8* pp. 26-31. https://doi.org/10.46501/IJMTST0801005

Article Info
Received: 24 November 2021; Accepted: 21 December 2021; Published: 31 December 2021

ABSTRACT

In recent times hand tracking and gesture tracking has become an integral part and it has opened up a variety of possibilities and challenges. There is rising interest in computer vision and therefore the fast development and improvement of accessible hardware which can support new developments in AI. In this article we outline, implement a handful of these possibilities and further summarize the challenges and future prospects in the spectrum of human user interaction and virtual reality. The motivation behind this paper is to reduce human interaction and dependency of devices to control the computer in view of COVID-19 spread. These results can drive further research and within the long run contribute to assist the use of virtual environments.

***KEYWORDS:** hand tracking, virtual reality, COVID-19, virtual environments*

1.INTRODUCTION

With the latest advances in virtual reality and its application in our daily lives, Bluetooth and wireless technology are becoming increasingly accessible. This paper proposes a visual AI system that uses computer vision to perform mouse, keyboard, and stylus functions using hand gestures and hand tip acquisition. Instead of using standard headsets or external devices, the suggested system traces the finger and hand movements to process the computer using a web camera or built-in camera. Because it is simple and effective, this solution may be continuously removed. the use of additional hardware, battery durability, and ultimately bring ease to the user.[1]

The AI mouse program is developed using the Python programming language, as well as OpenCV, a computer library. Models in the proposed visual AI mouse system use the Media Pipe package to track hands and title, as well as packages of Pynput, Autopy, PyGames, and PyAutoGUI to navigate the computer screen and perform tasks such as left-click, right-click, and then scroll. The findings of the proposed model show a very high level of accuracy, and the proposed model can work very well in real-world applications using only a CPU and no GPU.[2]

2.CHALLENGES

- Hand Detection: For ease of operation, it is important to facilitate detection and tracking of hands and isolate themselves from the visible universe. But dealing with material things can interfere with our ability to see and distinguish.[3] Due to differences in biological appearance and non-biological kinematics and / or factors such as temperature and texture may cause a problem.

- Hand Tracking: Hand tracking is reliable when the hands are in front of the face, perhaps to increase the visibility of individual viewing cameras. As the scope of VR activities expands to incorporate more natural everyday experiences where hands can participate in low-visibility activities, the limitations of tracking and visualization in this space region will probably become even more apparent.[4]
- Irregular Embodiment and Inclusivity: Involvement is a problem of increasingly important technologies, and the development of hand-tracking and demonstration presents a series of different challenges. An important part of hand tracking is to separate the skin from the surrounding area to form, and finally visualize, the strength of the hand. Another potential problem that has not been clearly considered is skin color and different skin tones.[5]

OPPORTUNITIES

- Increased Immersion: Immersion means the user's ability to perceive the physical environment using sensors and cameras found in the real world. Intensive research reveals that being able to see human hands being tracked in real time in a visible environment is a very attractive way to collaborate.[6][7] The supernatural knowledge of our hands in the physical world and their positions and functions, serves as a vivid reminder of our separation. Since our hands are in our direct view most of the time, removing them from the visual image may have different effects on our visual behavior.

- Effective Interaction: The main purpose of virtual reality is to provide the user with a natural way of communicating with the world created by simplicity. In its basic version, the user can explore the vast visible world by moving his head. The modern VR experience, on the other hand, almost always incorporates handmade inputs, from opening the doors to holding guns. [8] Manually accurate tracking can provide more accurate communication than would be possible with the controls, not only improving user immersion, but also the accuracy of their movements, which seems to be very important in the training context.

- Effective Communication: Through a series of gestures, human hands are used to express the words and sentences of the punctuation mark. Swipe and compression touch, for example, are used to fulfill orders.[9][10] During a casual conversation, however, informal gestures may seem important not only to the listener but also to the one to whom you are speaking. Indeed, recent research has suggested that people have difficulty distinguishing between the key sensations of the body without the hands. Given the extremely high precision of space and the space required to identify and track small exhibits, the benefits of physical touch on global communication are increasing. In the meantime, tracking and exposure of the hands, with their great and clear kinematics, should be a major factor in communication avatars.

4.GESTURE RECOGNITION BASED ON COMPUTER VISION

Touch-based touch perception is the most common current diagnostic method. Touch image information is collected by one or more cameras, and the collected data is processed in advance, including audio removal and information enhancement. [11]

Then, the separation algorithm is used to detect the targeted touch within the image. The separation of the current touch and the meaning can be obtained by processing and analyzing the video, and finally, the targeted touch is identified by the touch recognition algorithm. Touch-based touch detection is mainly made up of three parts: touch detection, touch analysis, and touch detection. The first step is to make the segments by tapping on the input image.

The process of touch separation mainly consists of two parts: the touch of the touch and the separation of the touch. The touch area process removes the touch circuit from the background complex in the self-tracking of the image that contains the touch and applies touch separation from the background. The touch phase segment separates the current touch from the background by using an algorithm after setting the touch. [12] Second, touch is organized and analyzed.

Currently, touch modeling techniques mainly include visual-based touch modeling and a 3D model-based touch model. The appearance-based touch model can be divided into vertical 2D models and moving models.

Finally, the touch recognition is based largely on conventional machine learning methods and sensory networks. There are many ways to identify vision-based touch using conventional machine learning models.

5.ALGORITHM USED

The various functions and conditions used in the system are explained in the flowchart of the real-time AI virtual system in Figure 1.

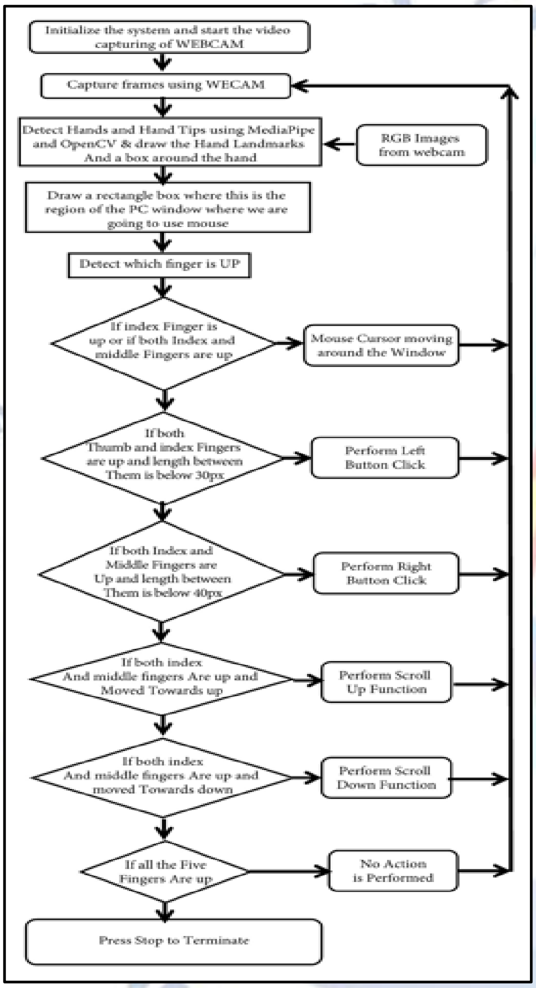

Fig. 1. Flowchart of the Real-time AI Virtual System.

- Camera Used in Visual AI System: The proposed AI mouse system is based on frames captured by a web camera on a portable computer or PC. Using the OpenCV Python computer library, a video recorder is created and the webcam will begin capturing video, as shown in Figure 2. The web camera captures and transmits frames to the visual AI system. [13]

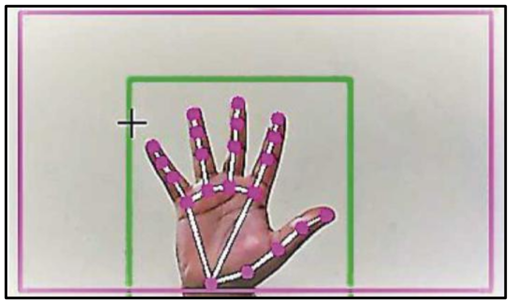

Fig. 2. Capturing Video using the Webcam.

- Capturing the Video and Processing: The visual AI mouse system uses a webcam where each frame is scanned until the program is completed. Video frames are processed from BGR to RGB color space to get hands on the video frame by frame as shown in the following code:

```
def findHands (self, img, draw = True):
imgRGB = cv2.cvtColor (img, cv2.COLOR_BG
R2RGB)
self. results = self. hands. Process (imgRGB) [14]
```

- Virtual Screen Matching: The visual AI mouse system uses a transformation algorithm, and converts finger links from the web camera screen to the full screen of a computer window to control the mouse. [15] When the hands are found and when we find out which finger is raised to perform a specific mouse function, a rectangular box is pulled towards the computer window in the web camera area where we travel across the window using the mouse cursor, as shown in the Figure. 3.

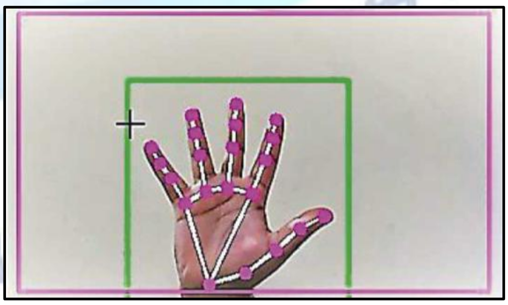

Fig. 3. Rectangular box for the area of the computer screen where we can move the cursor.

- Detecting Which Finger Is Up and Performing the Particular Mouse Function: In this section, we find out which finger at the top using the same fingerprint ID we found using MediaPipe and the corresponding upper finger link, as shown in Figure 4, and accordingly, a certain mouse function is performed.[16]

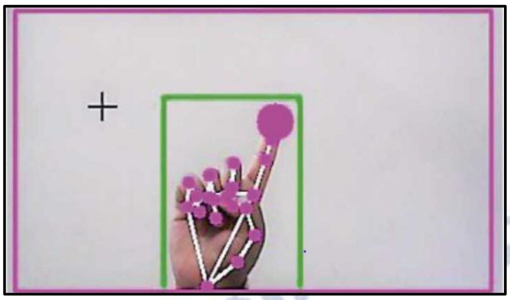

Fig. 4. Detection of which finger is up.

Mouse Functions Depending on the Hand Gestures and Hand Tip Detection Using Computer Vision:

- For the Mouse Cursor Moving around the Computer Window.
- For the Mouse to Perform Left Button Click.
- Assignable function for Fingers Touching
- Assignable function for Closed Hand
- For No Action to be Performed on the Screen.

6.EXPERIMENTAL RESULTS AND EVALUATION

In the proposed visual mouse AI system, the idea of improving human computer interaction using computer vision is given.

Contradictory comparisons of visual AI mouse system testing are difficult because only limited numbers of data sets are available. The test was performed 25 times by 4 people which resulted in 600 hand-labeled touches, and these tests were performed in different light conditions and at different distances from the screen, and each person tested the visual mouse system. AI 10 times in normal light conditions, 5 times in dim light, 5 times in the distance near the webcam, and 5 times in the distance from the webcam, and test results are tabulated in Table 1.[17]

Table. 1. Experimental Results

Hand tip gesture*	Mouse function performed	Success	Failure	Accuracy (%)
Tip ID 1 or both tip IDs 1 and 2 are up	Mouse movement	100	0	100
Tip IDs 0 and 1 are up and the distance between the fingers is <30	**Left button click**	99	1	99
Tip IDs 1 and 2 are up and the distance between the fingers is <40	**Right button click**	95	5	95
Tip IDs 1 and 2 are up and the distance between the fingers is >40 and both fingers are moved up the page	Scroll up function	100	0	100
Tip IDs 1 and 2 are up and the distancebetween the fingers is >40 and both fingers are moved down the page	Scroll down function	100	0	100
All five tip IDs 0, 1, 2, 3, and 4 are up	No action performed	100	0	100
Result		594	6	99

*Finger tip ID for respective fingers: tip Id 0: thumb finger; tip Id 1: index finger; tip Id 2: middle finger; tip Id 3: ring finger; tip Id 4: little finger.

From Table 1, it can be seen that the proposed visual AI mouse system achieved approximately 99% accuracy. From this 99% accuracy of the proposed visual AI mouse system, we come to know that the system works well. As can be seen in Table 1, the accuracy is under "Right-click" as this is the most difficult task for a computer to understand. The accuracy of the right click is low because the touch used to perform a particular mouse function is very difficult.[18] Also, the accuracy is very good and is high on all other touches. Compared to previous methods of visual mouse, our model performed very well with 99% accuracy. The accuracy graph is shown in Figure 5.

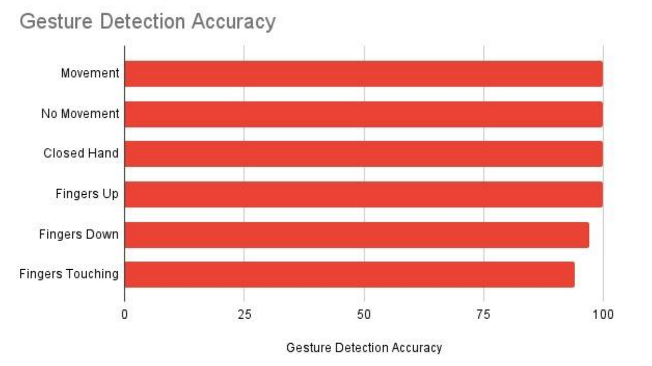

Fig. 5. Graph of Gesture Detection Accuracy.

Table 2 shows a comparison between the existing models and the proposed AI virtual model in terms of accuracy.

Table. 2. Comparison with Existing Models.

Existing models	Accuracy (%)
Virtual mouse system using RGB-D images and fingertip detection [16]	96.13
Palm and finger recognition based [17]	78
Hand gesture-based virtual mouse [18]	78
The proposed AI virtual mouse system	99

As can be seen in Table 2, the suggested artificial intelligence virtual mouse and keyboard outperformed the other virtual models when it came to accuracy when compared to their counterparts. [19] With fingertip detection, the proposed model is unique in that it can perform the majority of mouse functions such as left and right clicks as well as scrolling up and down as well as mouse cursor movement, and that it can also be used to control the PC just like a physical mouse but in a virtual environment. Figure 6 depicts a graph that compares the two models side by side.

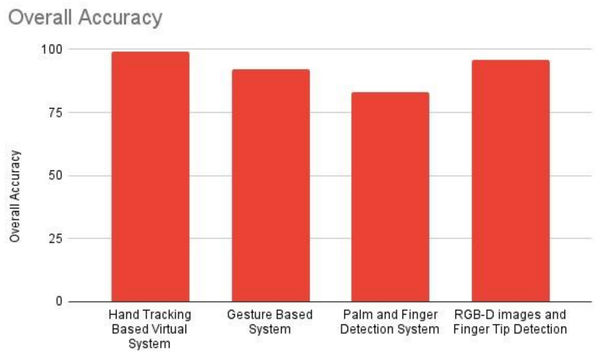

Fig. 6. Graph for Comparison between the Models.

7. OTHER ALTERNATIVES

- ClayAIR: Its hand tracking solutions aim at higher performance, quicker implementation time and higher accuracy. It can predict 22 3D key point coordinates. Using regression algorithms trained on 1.4 million images including real and synthetic images. It is being used by some leading tech giants like enovo, NReal, QUalcomm to bring virtual reality to the digital world.

- SOTA Hardware :
  Data Gloves: Data gloves are pure VR devices in the sense that it can detect activity of the joints and on the other hand the feedback enables the user to feel the virtual targets in a pseudo-physical sense. They are especially famous in the VR field since they are highly accurate and the inference time is less. Additionally, they are a great way to collect data of hand-landmarks for machine learning models. But since photoelectric sensors and position trackers are costly, the production and maintenance of these gloves is also high.
- Inertial Sensors : The Nintendo Wii was the commercial release of inertial sensors. They are composed of an actuator and a sensor which help to collect and obtain information about gestures. Built with an accelerometer and an infrared sensor, they can capture the user's wrist and arm gestures.
- KCF : KCF algorithm or Kernelized correlation filter algorithm is mainly focused around creating large number.of training examples by shifting the target area in a circular manner. It was widely used for object tracking and is the base of many recent tracking algorithms. Unfortunately the algorithm doesn't perform well in case of scale variations i.e changes in the size of target objects. Additionally it is not easy to train it for detecting multiple landmarks.

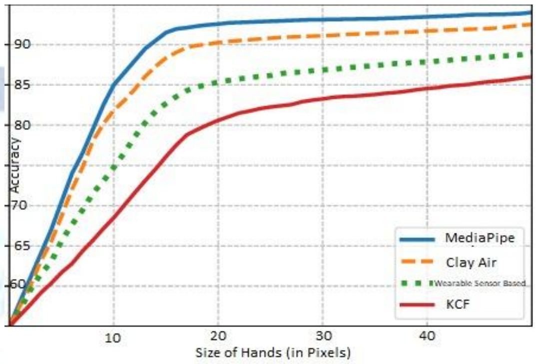

Fig. 7. Comparison Graph.

8.APPLICATIONS

- The suggested model has a better accuracy of 95 percent, which is much higher than the accuracy of other proposed models for virtual systems, and it has a wide range of applications in many fields.
- Since it is not safe to use the devices by touching them during the COVID-19 condition due to the possibility of the virus spreading by touching the devices, it is recommended that an artificial intelligence virtual system be used to operate the PC mouse functions without utilising physical devices.[20]
- With no need for external devices, the system may be utilised to control robots and automation systems directly.
- Drawing 2D and 3D pictures using the AI virtual system is possible via the use of hand motions.
- Playing virtual reality and augmented reality games without the usage of a wireless or wired mouse device is possible with this gadget.
- People who suffer from hand illnesses can utilise this technology to manage the mouse functionalities of a computer with ease.[21]
- In the realm of robotics, the suggested system, such as HCI, may be utilised to control robots and other machinery.
- In the fields of design and architecture, the suggested technology may be utilised for virtual prototyping and design in general.[22]

9.CONCLUSION

The primary goal of the visual AI mouse system is to allow users to control mouse cursor functions with a hand touch rather than by manipulating things. Hand gestures and hand tips are detected and processed by the proposed system, which may be accessible by a webcam or a built-in camera that can be used to access the system.

Based on the findings of the model, we can infer that the suggested artificial intelligence system has done extremely well and is extremely accurate when compared to current models, and that the model addresses the majority of the constraints of the existing system.

Being very precise, the suggested model may be utilised for real-world applications as well. For example, it can be used to decrease the spread of COVID-19 and can eliminate the need for portable equipment.

Some drawbacks of the model include a modest loss in the precision of the right-click function and the difficulty of selecting text by clicking and dragging. As a result, we will seek to address these restrictions by designing a fingerprint acquisition method that will yield more accurate findings in the future.

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-based-interaction-for-the-nextcomputing-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., Michelakakis, 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 InfoCommunications," 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] L. Thomas, "Virtual mouse using hand gesture," International Research Journal of Engineering and Technology (IRJET, vol. 5, no. 4, 2018. View at: Google Scholar.

[15] P. Nandhini, J. Jaya, and J. George, "Computer vision system for food quality evaluation—a review," in Proceedings of the 2013 International Conference on Current Trends in Engineering and Technology (ICCTET), pp. 85–87, Coimbatore, India, July 2013. View at: Publisher Site | Google Scholar.

[16] J. Jaya and K. Thanushkodi, "Implementation of certain system for medical image diagnosis," European Journal of Scientific Research, vol. 53, no. 4, pp. 561–567, 2011. View at: Google Scholar.

[17] P. Nandhini and J. Jaya, "Image segmentation for food quality evaluation using computer vision system," International Journal of Engineering Research and Applications, vol. 4, no. 2, pp. 1–3, 2014. View at: Google Scholar.

[18] 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.

[19] J. T. Camillo Lugaresi, "Media Pipe: A Framework for Building Perception Pipelines," 2019, https://arxiv.org/abs/1906.08172. View at: Google Scholar.

[20] Google, MP, https://ai.googleblog.com/2019/08/on-device-real-time-hand-tracking-with.html.

[21] V. Bazarevsky and G. R. Fan Zhang. On-Device, Media Pipe for Real-Time Hand Tracking.

[22] K. Pulli, A. Baksheev, K. Kornyakov, and V. Eruhimov, "Realtime computer vision with openCV," Queue, vol. 10, no. 4, pp. 40–56, 2012. View at: Publisher Site | Google Scholar.